Regulation of TGFβ signaling by long non-coding RNAs and ADP-ribosylation

PANAGIOTIS PAPOUTSOGLOU
Abstract


Transforming growth factor β (TGFβ) signaling pathways participate in embryonic development and tissue homeostasis and have a dual role in cancer. TGFβ acts as a tumor suppressor that promotes cell cycle arrest and apoptosis at initial stages of tumorigenesis. In contrast, TGFβ induces epithelial to mesenchymal transition (EMT), a normal embryonic process which is employed by advanced cancers, in order to acquire mesenchymal traits and metastasize.

Bone morphogenetic protein (BMP) family members belong to the TGFβ superfamily and are involved in cell differentiation, development and bone formation.

Non-coding RNAs (ncRNAs) are not translated into proteins, are important regulators of gene expression and physiological processes and are often de-regulated in cancer. They control gene expression through physical association with chromatin, DNA, RNA molecules or proteins.

Poly(ADP-ribose) polymerases (PARPs) catalyze the poly (ADP)-ribosylation of proteins, whereas the enzyme poly(ADP-ribose) glycohydrolase (PARG) removes ADP-ribose units. Members of the PARP family function in the DNA damage response, regulation of transcription and cell death.

In this thesis, we investigated the importance of the TGFβ signaling pathway in regulating the expression of long non-coding RNAs (lncRNAs). We identified TGFβ-regulated lncRNAs and observed that a substantial number of them act in a feedback loop to modulate the magnitude of TGFβ signaling. Interestingly, the nuclear lncRNA TGFB2-antisense RNA 1 (TGFB2-AS1) is induced by TGFβ and negatively regulates expression of members of the TGFβ and BMP pathways, through interaction with EED, a protein of the polycomb repressor complex 2 (PRC2). Also, TGFβ signaling promoted the expression of mir-100-let-7a-2-mir-125b-1 cluster host gene (MIR100HG), which enhanced TGFβ signaling and affected TGFβ-mediated cell cycle arrest. The MIR100HG-derived miRNAs let-7a-2-3p, miR125b-5p and miR-125b-1-3p, were also induced by TGFβ. In contrast, the long intergenic non-protein coding RNA 707 (LINC00707), was reduced in response to TGFβ and affected the expression of a group of genes related to inflammatory responses and interferon-γ (IFN-γ) signaling.

We also report that TGFβ and BMP pathways are regulated by ADP-ribosylation of Smad proteins, the signaling mediators of these pathways. We observed that PARP1 and PARP2 attenuated, while PARG favored TGFβ signaling. Furthermore, PARP1 negatively regulated BMP signaling, by ADP-ribosylating Smad1 and Smad5, whereas PARG enhanced BMP signaling by de-ADP-ribosylating Smads.

Collectively, we provide evidence that lncRNAs and ADP-ribosylating enzymes modulate TGFβ and BMP signaling pathways and propose models for their molecular mechanisms and functional roles.

Keywords: TGFβ, signal transduction, long non-coding RNAs, ADP-ribosylation, transcriptional regulation, chromatin remodeling

Panagiotis Papoutsoglou, Science for Life Laboratory, SciLifeLab, Box 256, Uppsala University, SE-75105 Uppsala, Sweden.

© Panagiotis Papoutsoglou 2018

ISSN 1651-6206
urn:nbn:se:uu:diva-364107 (http://urn.kb.se/resolve?urn=nbn:se:uu:diva-364107)
To my parents
This thesis is based on the following papers, which are referred to in the text by their Roman numerals.


* Authors contributed equally to this work

Reprints were made with permission from the respective publishers.
## Contents

### Introduction ............................................................................................. 13

1. Transforming growth factor β (TGFβ) signaling pathway ............. 13  
   1.1 Canonical TGFβ signaling .......................................................... 13  
   1.2 The Smad family ...................................................................... 14  
   1.3 Protein kinase signaling downstream of TGFβ ......................... 15  
   1.4 TGFβ-mediated physiological responses ................................. 15  
   1.5 TGFβ in cancer........................................................................... 18  

1.6 Bone morphogenetic protein (BMP) signaling ............................. 20  
   1.7 BMP-mediated physiological responses .................................. 21  

2. Long non-coding RNAs ...................................................................... 23  
   2.1 Non-coding RNAs ........................................................................ 23  
   2.2 Classification of IncRNAs ............................................................ 24  
   2.3 Biological functions of IncRNAs .................................................. 24  
   2.4 Molecular functions of IncRNAs .................................................. 26  
      2.4.1 Nuclear IncRNAs can function in cis or in trans ............... 26  
      2.4.2 IncRNAs are involved in transcriptional and epigenetic regulation of gene expression ......................................................... 27  
      2.4.3 IncRNAs mediate post-transcriptional regulation ............... 28  
      2.4.4 Cytoplasmic IncRNAs ........................................................... 29  
   2.5 Polycomb repressor complex 2 (PRC2) and IncRNAs interplay .. 29  
   2.6 TGFβ-regulated IncRNAs ............................................................. 32  
   2.7 IncRNAs as regulators of the TGFβ pathway .............................. 34  

3. MicroRNAs ......................................................................................... 36  
   3.1 Biogenesis and functions of microRNAs................................. 36  
   3.2 Micro-RNAs in TGFβ signaling .................................................. 37  

4. The poly(ADP)ribose polymerase family ........................................... 38  
   4.1 Poly(ADP)ribosylation ................................................................. 38  
   4.2 Poly(ADP)ribose polymerases (PARPs) ...................................... 38  
   4.3 Catabolism of Poly(ADP) ribosylation ...................................... 39  
   4.4 Cellular functions of PARPs ......................................................... 40  
   4.5 PARPs in TGFβ signaling .......................................................... 41
5. Present investigations

**Paper I:** The *TGFB2-AS1* lncRNA regulates TGFβ signaling by modulating corepressor activity.

**Paper II:** TGFβ signaling down-regulates *LINC00707* to inhibit inflammatory responses.

**Paper III:** The non-coding *MIR100HG* RNA mediates cytostatic responses of epithelial cells to transforming growth factor β.

**Paper IV:** Fine-tuning of Smad protein function by poly(ADP-ribose) polymerases and poly(ADP-ribose) glycohydrolase during transforming growth factor β signaling.

**Paper V:** Regulation of bone morphogenetic protein signaling by ADP-ribosylation.

6. Future perspectives

**Paper I:** The *TGFB2-AS1* lncRNA regulates TGFβ signaling by modulating corepressor activity.

**Paper II:** TGFβ signaling down-regulates *LINC00707* to inhibit inflammatory responses.

**Paper III:** The non-coding *MIR100HG* RNA mediates cytostatic responses of epithelial cells to transforming growth factor β.

**Paper IV:** Fine-tuning of Smad protein function by poly(ADP-ribose) polymerases and poly(ADP-ribose) glycohydrolase during transforming growth factor β signaling.

**Paper V:** Regulation of bone morphogenetic protein signaling by ADP-ribosylation.

Acknowledgments

References
Abbreviations

α-ESA: α-eleostearic acid
α-SMA: α-smooth muscle actin
AIF: apoptosis-inducing factor
ANRIL: antisense ncRNA in the INK4 locus
ARH3: ADP-ribosyl hydrolase 3
ARTS: apoptosis-related protein in the TGFβ signaling pathway
asFGFR2: antisense RNA from FGFR2
BC1: brain cytoplasmic RNA 1
BER: base excision repair
BMP: bone morphogenetic protein
BMPRI: bone morphogenetic protein receptor type I
BMPRII: bone morphogenetic protein receptor type II
BREs: BMP-responsive elements
CDK: cyclin-dependent kinase
CDKI: cyclin-dependent kinase inhibitor
ceRNA: competing endogenous RNA
circRNA: circular RNA
co-Smad: common mediator Smad
DAPK: death-associated protein kinase
DEANR1: definitive endoderm-associated IncRNA1
DSBs: double strand breaks
EED: embryonic ectoderm development
EMT: epithelial to mesenchymal transition
ERK: extracellular-signal regulated kinase
eRNA: enhancer RNA
EZH2: methyltransferase enhancer of zeste 2
FGFR2: fibroblast growth factor receptor 2
FN1: fibronectin 1
FOXA2: forkhead box A2
GADD45β: growth arrest and DNA damage-inducible 45β
GAS5: growth arrest-specific 5
GBM: glioblastoma multiforme
GSK-3: glycogen synthase kinase-3
HCC: hepatocellular carcinoma
HOTAIR: HOX transcript antisense RNA
IFN-γ: interferon-γ
IL-2: interleukin-2
IL-4: interleukin-4
IL-11: interleukin-11
I-Smads: inhibitory Smads
JNK: Jun N-terminal kinase
KCNQ1OT1: KCNQ1 opposite strand or antisense transcript 1
lnc-LFAR1: fibrosis-associated IncRNA1
lncRNA: long non coding RNA
lncRNA-ATB: IncRNA-activated by TGFβ
lncRNA-LINP1: IncRNA in nonhomologous end joining pathway 1
LINC00707: long intergenic non-protein coding RNA 707
MEG3: maternally expressed gene 3
MET: mesenchymal to epithelial transition
MH1: Mad homology 1
MH2: Mad homology 2
MIR100HG: mir-100-let-7a-2-mir-125b-1 cluster host gene
miRNA: microRNA
miRISC: miRNA-induced silencing complex
mRNA: messenger RNAs
NAD+: nicotinamide adenine dinucleotide
NER: nucleotide excision repair
ORF: open reading frame
PARs: poly(ADP)ribose polymers
PARylation: poly(ADP)ribosylation
PARG: poly(ADP-ribose) glycohydrolase
PARP: poly(ADP-ribose) polymerase
PDCD4: programmed cell death 4
PI3K: phosphatidylinositol 3’-kinase
piRNAs: PIWI-interacting RNAs
PRC2: polycomb repressive complex 2
pre-miRNA: precursor miRNA
pri-miRNA: primary miRNA
Rb: Retinoblastoma protein
RNA-FISH: RNA-fluorescence in situ hybridization
rSBEs: RNA Smad-binding elements
R-Smads: receptor-activated Smads
RT-qPCR: real-time quantitative polymerase chain reaction
siRNA: small interfering RNA
Smad: Small mothers against decapentaplegic
TAK1: TGFβ-activated kinase 1
TARG1: terminal ADP-ribose protein glycohydrolase 1
TGFβ: transforming growth factor β
TGFB2-AS1: TGFB2 antisense RNA 1
Th: T helper
TIEG1: TGF-β-inducible early response gene-1
TLINC: TGFβ-induced long noncoding RNA
TβRI: transforming growth factor β receptor type I
TβRII: transforming growth factor β receptor type II
TRAF6: tumor necrosis factor α receptor associated factor 6
Tsg: Twisted Gastrulation
UCA1: urothelial cancer associated 1
Uchl1: ubiquitin carboxy-terminal hydrolase L1
VEGF: vascular endothelial growth factor
VIM: vimentin
XCI: X chromosome inactivation
YY1: Yin Yang
4EBP1: 4E-binding protein 1
Introduction

1. Transforming growth factor β (TGFβ) signaling pathway

1.1 Canonical TGFβ signaling

Signaling pathways regulate several vital cellular processes, such as cell communication, development, regulation of gene expression in response to micro-environmental changes, immune responses and proliferation. Among different signaling pathways, TGFβ signaling regulates embryonic development and adult tissue homeostasis, under physiological conditions. In addition, a rather complicated role has been attributed to TGFβ signaling during cancer progression. In primary tumors TGFβ acts as a tumor suppressor by promoting cell cycle arrest and/or apoptosis. On the other hand, in advanced malignancies it promotes migration and metastasis (Heldin et al., 2009).

The TGFβ pathway is activated when extracellular TGFβ ligands bind to type I and type II TGFβ receptors (TβRI and TβRII) which possess serine/threonine kinase and much weaker tyrosine kinase activity in their intracellular domains. TβRII receptor interacts with high affinity with TGFβ ligands and a heterotetrameric complex consisting of two TβRI and two TβRII is formed. This event allows the phosphorylation of TβRI by TβRII, a modification which enforces TβRI to phosphorylate the transcription factors, that belong to the small mothers against decapentaplegic (Smad) family, Smad2 and Smad3 (receptor-activated Smads, R-Smads) at their C-terminal regions (Ikushima and Miyazono, 2010). Phosphorylated Smad2 and Smad3 interact with Smad4 and the resulting complex, in co-operation with co-transcription factors and chromatin modifying enzymes, positively or negatively regulates gene expression (Figure 1). TGFβ signaling is subjected to negative regulation, mediated mainly by Smad7, which antagonizes Smad2 and Smad3 for binding to TβRI or, alternatively, promotes the ubiquitination and subsequent degradation of TβRI (Lönn et al., 2009). Additional mechanisms of negative regulation of TGFβ signaling include the SnoN and Ski proteins, which are mainly nuclear proteins that bind to R-Smads and Smad4 and prevent the formation of active trimeric Smad complexes (Deheuninck and Luo, 2009). The inhibitory effects of Smad7, SnoN and Ski on TGFβ signaling are abolished by the E3-ubiquitin ligase Arkadia, which catalyzes the ubiquitination of these proteins, ultimately leading to their degradation (Koinuma et al., 2003; Nagano et al., 2007).
1.2 The Smad family

The Smad family consists of R-Smads, common mediator (Co-Smad) and inhibitory Smads (I-Smads). The first subgroup involves the TGFβ-activated Smad2 and Smad3, as well as the Smad1, Smad5 and Smad8, which are activated in response to bone morphogenetic protein (BMP) pathway. The co-Smad (Smad4) forms complexes with R-Smads, upon activation of TGFβ signaling and it is required for the Smad-dependent transcriptional regulation. The I-Smads, Smad6 and Smad7, restrict TGFβ (Smad7) and BMP (both Smad6 and Smad7) signaling. Interestingly, TGFβ induces *SMAD7* mRNA expression (Nakao et al., 1997) and BMP up-regulates *SMAD6* and *SMAD7* (Takase et al., 1998), in order to create negative feedback loops that limit the activity of these pathways.

The R-Smad proteins are widely conserved and contain different domains, which mediate distinct functions. Structurally, R-Smads can be divided in three main parts, i.e. an amino-terminal Mad homology 1 (MH1) domain, a central linker and a carboxy-terminal Mad homology 2 (MH2) domain. The MH1 domains of Smad1, Smad3, Smad4, Smad5 and Smad8 bind DNA. The MH2 domain mediates interaction with a variety of proteins, such as TβRI, co-transcription factors, chromatin remodeling factors and SMAD4 (Macias et al., 2015). The linker domain serves as a platform for post-translational modifications, which affect the activity and stability of Smads, including phosphorylation by several kinases, such as mitogen-activated protein kinases (MAPKs), which include the extracellular-signal regulated kinase (ERK), Jun N-terminal kinase (JNK) and p38 kinase, the tyrosine kinase Src, phosphatidylinositol 3’-kinase (PI3K), cyclin-dependent kinases (CDKs), rho-associated protein kinase, calcium calmodulin-dependent kinase and glycogen synthase kinase-3 (GSK-3) (Kamato et al., 2013). In addition, E3 ubiquitin ligases bind to Smads, via the linker domain and promote their proteasome-dependent degradation (Imamura et al., 2013).

Recently, Smads have been shown to have RNA-binding properties. For example, Smad3 associates with the long non-coding RNAs (lncRNAs) *growth arrest-specific 5* (*GAS5*) (Tang et al., 2017) and *brain cytoplasmic RNA 1 (BC1)* (Wang et al., 2018c), via RNA Smad-binding elements (rSBEs), which reside in the sequence of these transcripts. These interactions affect the ability of Smad3 to bind to promoters of target genes or its subcellular localization. In addition, the *definitive endoderm-associated lncRNA 1 (DEANR1)* directly associates with and recruits Smad2/3 to the *forkhead box A2* (*FOXA2*) promoter, thus promoting human endoderm differentiation (Jiang et al., 2015). Moreover, the TGFβ/Smad3-interacting long non-coding RNA (*Inc-TSI*) binds to the MH2 domain of Smad3 and prevents its interaction with the TβRI, thereby limiting TGFβ-induced renal fibrogenesis (Wang et al., 2018b). Interestingly, the TGFβ-activated R-Smads (Smad2/3) and the BMP-activated Smad1/5 directly interact with a subset of primary miRNAs (pri-miRNAs) and
facilitate their maturation, in a ligand-dependent manner (Davis et al., 2008). The specificity of the Smad-pri-miRNA interaction is provided by rSBEs, which are located in the pri-miRNA sequence (Davis et al., 2010). Thus, R-Smads are not only considered as classical DNA-binding transcription factors, but also as molecular partners of non-coding RNAs.

1.3 Protein kinase signaling downstream of TGFβ

In addition to Smad-mediated TGFβ signaling, other pathways which are not dependent on Smad protein activation are also induced upon binding of TGFβ ligands to their receptors. Non-Smad signaling includes Erk, JNK, p38 kinase pathways, and phosphatidylinositol-3 kinase (PI3K)/Akt. Furthermore, the ubiquitin ligase tumor necrosis factor α receptor associated factor 6 (TRAF6) binds to TβRI and promotes the cleavage of the receptor, whose intracellular domain is transferred into the nucleus and affects genes involved in invasiveness. TRAF6 and its relative TRAF4 bind to TβRI and activate, via lysine 63-linked poly-ubiquitination, the TGFβ-activated kinase 1 (TAK1), which acts as an upstream activator for many other protein kinases of the MAPK family (Massagué, 2012). TGFβ can activate the PI3K/Akt pathway in a TRAF6-dependent manner, whereby TRAF6 poly-ubiquitininatates the p85α subunit, leading to the complex formation between p85α and TβRI and the subsequent activation of the PI3K/Akt pathway (Hamidi et al., 2017). Moreover, TβRII directly phosphorylates Par6, a member of the polarity complex which is important in supporting the apical/basal polarity of epithelial cells, thereby destabilizing this complex (Massagué, 2012). Similarly to canonical TGFβ/Smad pathways, the Smad-independent pathways also lead to growth arrest, induction of apoptosis and activation of the epithelial to mesenchymal transition (EMT) program.

1.4 TGFβ-mediated physiological responses

1.4.1 Cell growth arrest

TGFβ has tumor suppressive properties at initial phases of tumor progression. These properties are attributed, in part, to the TGFβ-mediated induction of cytostasis in normal or cancer cells. TGFβ orchestrates cell cycle arrest, through induction of anti-proliferative factors and suppression of pro-proliferative genes. For example, TGFβ up-regulates the expression of the cyclin-dependent kinase inhibitors (CDKIs) p21 (encoded by the CDKN1A gene), p15 (encoded by the CDKN2B gene) and p27 (encoded by the CDKN1B gene) (Heldin and Moustakis, 2012). In contrast, the proto-oncogene c-Myc is repressed by TGFβ at the transcriptional level (Frederick et al., 2004). The inhibition of c-Myc expression by TGFβ involves the formation of a transcrip-
tion repressor complex, consisting of Smad3, E2F4/5 and p107, at the promoter of c-Myc. Moreover, TGFβ down-regulates members of the inhibitor of differentiation (Id) family, which encodes transcription factors that promote cell proliferation (Damdinsuren et al., 2006). Another mechanism of cell growth arrest includes the repression of the CDK tyrosine phosphatase Cdc25A by TGFβ, which dephosphorylates CDKs and promotes cell cycle progression in proliferating cells (Iavarone and Massagué, 1999). In addition, the cell growth inhibitor and translation initiation factor 4E-binding protein 1 (4EBP1) halts proliferation in response to TGFβ (Azar et al., 2009). Interestingly, IncRNAs have also been implicated in the TGFβ-induced cell cycle arrest. The antisense non-coding RNA in the INK4 locus (ANRIL), an lncRNA expressed within the same genetic locus as p14, p15 and p16, has been shown to inhibit CDKN2B and TGFB1 mRNA expression. Cell proliferation is enhanced in human esophageal squamous cell carcinoma, as a consequence of TGFβ1 and p15 repression by ANRIL (Chen et al., 2014).

1.4.2 Apoptosis
TGFβ signaling exerts its anti-tumor functions not only by inhibiting cell proliferation, but also by promoting apoptosis, or programmed cell death. TGFβ promotes apoptosis in some cancer types, such as hepatocellular carcinoma, through induction of pro-apoptotic genes. Some of the TGFβ-regulated pro-apoptotic genes are the death-associated protein kinase (DAPK), the growth arrest and DNA damage-inducible 45β (GADD45β), the TGF-β-inducible early response gene-1 (TIEG1) (Pardali and Moustakas, 2007) and the programmed cell death 4 (PDCD4) (Zhang et al., 2006). In addition, downstream components of the apoptotic machinery, such as Daxx and apoptosis-related protein in the TGFβ signaling pathway (ARTS) positively influence TGFβ-induced apoptosis. Although TGFβ signaling is mainly pro-apoptotic, it can also be anti-apoptotic in certain cell types. For instance, TGFβ mediates pro-survival signals, through activation of PI3K-Akt signaling in NMuMG and 4T1 mouse mammary epithelial cells (Shin et al., 2001). In addition, TGFβ induces the expression of IncRNA-Smad7, a lncRNA, which is characterized by anti-apoptotic properties, in NMuMG cells. Silencing IncRNA-Smad7 compromises TGFβ-induced survival in NMuMG cells (Arase et al., 2014). Overall, the positive or negative effect of TGFβ signaling on apoptosis is highly context-dependent.

1.4.3 Epithelial to mesenchymal transition (EMT)
Epithelial to mesenchymal transition is a developmental process, whereby epithelial cells lose their contacts with neighboring cells, their polarity and acquire mesenchymal properties, which allow them to migrate far from their initial location (Nieto et al., 2016). The induction of EMT is important for embryogenesis and organ development, but it can also contribute to pathological disorders, such as metastasis and fibrosis (Nieto et al., 2016). The process of
EMT is initiated by several input signals. Among them, TGFβ signaling is a potent inducer of EMT (Moustakas and Heldin, 2016). TGFβ orchestrates a pro-EMT program, by inducing the expression of mesenchymal markers, such as *fibronectin 1* (*FN1*) and *vimentin* (*VIM*), and transcription factors, which facilitate the establishment of EMT, such as Snail, Slug, ZEB1 and ZEB2 (Moustakas and Heldin, 2016). Concomitantly, TGFβ represses the expression of epithelial markers, such as E-cadherin and ZO-1 and promotes the rearrangement of the actin cytoskeleton to form stress fibres. Furthermore, TGFβ regulates the expression of specific miRNAs that affect the EMT. For example, TGFβ reduces the expression of the miR-200 family members, which antagonize EMT by targeting ZEB1 and ZEB2. Interestingly, ZEB proteins repress miR-200 expression at the transcriptional level, in order to potentiate the EMT program, leading to the formation of double negative feedback loops (Bracken et al., 2008).

TGFβ also controls the expression of several lncRNA genes which ultimately contribute to the process of EMT. Some examples include the long non-coding RNA activated by TGFβ (*lncRNA-ATB*), which is induced by TGFβ and promotes EMT through up-regulation of ZEB1 and ZEB2, the HOXA transcript induced by TGFβ (*lncRNA-HIT*), which favors EMT, through targeting E-cadherin, the lncRNA Inc-Spry1, which is down-regulated by TGFβ and inhibits EMT and the *lncRNA in nonhomologous end joining pathway 1* (*lncRNA-LINP1*), a TGFβ-repressed lncRNA, which has an inhibitory role in EMT and cell invasion in lung cancer (Zhang et al., 2018a). In addition, TGFβ promotes the alternative splicing of the pre-RNA, which encodes the protein phosphatase 1 regulatory subunit 10 (PPP1R10, also designated as *PNUTS*), in order to produce a lncRNA that favors EMT by sponging the miR-205 and thus stabilizing ZEB1 protein levels (Grelet et al., 2017).

**1.4.4 Immune response**

TGFβ is an important regulator of immune responses, as it is involved in the differentiation of immune cells, as well as in the modulation of inflammatory responses (Travis and Sheppard, 2014). TGFβ is a potent suppressor of immune responses, a function which is exerted in multiple ways. First, TGFβ inhibits the differentiation of effector T helper (Th) cells (Dahmani and Delisle, 2018). In addition, it promotes the differentiation of CD4+ T cells to regulatory T cells, by inducing the transcription factor FOXP3 (Chen and Konkel, 2015). Moreover, TGFβ halts proliferation of B cells and T cells, by up-regulating cell cycle inhibitors (p15, p21, p27) and down-regulating factors that mediate cell cycle progression (c-myc, cyclins) (Pardali and Moustakas, 2007). In a different mechanism, TGFβ lowers the expression of inflammatory cytokines, such as interferon-γ (*IFN-γ*), interleukin-2 (*IL-2*) and interleukin-4 (*IL-4*), in a Smad2/3-dependent manner (Yoshimura et al., 2010). The TGFβ-mediated *IFN-γ* inhibition is achieved due to the repression of the transcription factor T-bet, a major determinant of *IFN-γ* expression (Neurath et al., 2002).
Finally, TGFβ suppresses different types of immune cells, such as natural killer cells, dendritic cells and macrophages (Yoshimura et al., 2010). On the other hand, TGFβ promotes the differentiation of another type of T helper cells, the Th17 cells, which express high levels of IL-17 and target bacterial and fungal pathogens. This type of Th cells exhibits pro-inflammatory properties and is implicated in autoimmune diseases (David and Massagué, 2018). Overall, the suppressive effects of TGFβ on immune responses may have protective roles against tumor-promoting chronic inflammation. However, the inhibitory effects of TGFβ on immune cells may promote tumorigenesis evasion of immune surveillance.

![Image](image_url)

*Figure 1. Illustration of the Smad-dependent TGFβ pathway and the TGFβ-regulated biological processes.*

### 1.5 TGFβ in cancer
Cancer is a complex disease, which is characterized by genetic or epigenetic alterations that lead to aberrant cell proliferation and growth. The mutations in the genome, which ultimately lead to the onset of cancer are a consequence of the genetic background or environmental factors (Migliore and Coppedè, 2002). Normal cells can be transformed to cancer cells, which will then give
rise to malignant tumors, if they acquire some or all of the following properties: sustaining proliferative signaling, evading growth suppressors, resistance to cell death, replicative immortality, induction of angiogenesis and activation of invasion and metastasis (Hanahan and Weinberg, 2011). Moreover, increased genomic instability, reprogramming of cell metabolism, inducing tumor-promoting inflammation and escaping immune destruction are additional traits of cancer cells (Hanahan and Weinberg, 2011). As discussed above, TGFβ signaling favors cytostatic or pro-apoptotic effects in normal cells and in cancers of epithelial origin. In fact, many tumors acquire mutations in core components of the TGFβ pathway, such as the genes encoding TGFβ receptors and Smads, in order to bypass the tumor-suppressing effects of TGFβ (Abraham et al., 2018). Alternatively, tumors can maintain intact core signaling, but acquire mutations in downstream components, that mediate the tumor suppressor effects of TGFβ, i.e. genes encoding cell cycle inhibitors or pro-apoptotic genes. In the latter case, advanced tumors, by preserving a functional TGFβ pathway, can hijack the tumor-promoting effects of the pathway (EMT, evasion of immunosurveillance), thereby converting TGFβ from a guardian that prevents against tumor progression, to a promoter of tumor invasion and metastasis (Massagué, 2008). Other tumor promoting properties of TGFβ include the induction of angiogenesis, through the up-regulation of the vascular endothelial growth factor (VEGF) (Haque and Morris, 2017). This is an important step for providing the tumor with oxygen and nutrients. The induction of glycolysis (Guido et al., 2012; Rodríguez-García et al., 2017) is a process that gives a growth advantage to tumors that expand under low oxygen levels. Overall, TGFβ plays a dual role in tumor progression and affects either positively or negatively, several hallmarks of cancer (Figure 2).
Figure 2. Crosstalk between TGFβ signaling and several hallmarks of cancer. The effects of TGFβ on cancer progression can be divided into two main branches. In the tumor-suppressing branch, TGFβ exhibits anti-tumorigenic functions by restricting cell proliferation, survival and pro-tumorigenic chronic inflammation. In contrast, the tumor-promoting branch involves biological processes that contribute to the induction of angiogenesis, the metabolic switch towards the glycolytic pathway, the escape of elimination of cancer cells by the immune system and the activation of invasion and metastasis.

1.6 Bone morphogenetic protein (BMP) signaling

Bone morphogenetic protein ligands are part of the TGFβ superfamily. The about 15 BMP ligands bind to type I and type II BMP receptors (BMPRI and BMPRII) (Yadin et al., 2016). Analogous to TGFβ signaling, BMPRI is phosphorylated by BMPRII, and the effector R-Smads (Smad1, Smad5 and Smad8) are phosphorylated by BMPRI. The phosphorylated Smads interact with Smad4 and the complex enters into the nucleus to regulate the expression of BMP-target genes (Davis et al., 2016). This regulation is achieved by binding of Smads to specific genomic loci at the promoters of target genes denoted BMP-responsive elements (BREs) and is coordinated by co-transcription factors, which further determine the specificity of Smad-binding sites at the promoter regions. BMP signaling activates the transcription of members of the Id family (Hollnagel et al., 1999). In addition, it induces the expression of the I-Smad genes SMAD6 and SMAD7 (Takase et al., 1998). Similar to TGFβ receptors, the BMP receptors also activate protein kinases, such as TAK1 and
downstream MAPKs, which mediate their diverse physiological effects (Monzen et al., 1999).

Like other pathways, BMP signaling is subjected to fine-tuning so that the transcriptional responses mediated by this pathway are tightly controlled. Thus, negative regulation of BMP signaling is important and includes antagonists of BMP ligands, co-receptors and different intracellular regulators, including the I-Smads. The antagonists of BMP ligands include the Chordin/Noggin, Twisted Gastrulation (Tsg) and DAN/Cerberus families, the latter including Gremlin1/GREM1; they mainly act by masking the BMP receptor binding regions of the ligands, preventing them from interacting with their receptors (Sieber et al., 2009). Recently, it has been observed that BMP signaling modulates the expression of genes that give rise to IncRNAs. One such example is the IncRNA *urothelial cancer associated 1* (*UCA1*), which is induced by BMP9 stimulation in bladder cancer cells (Gou et al., 2018). However, the possible involvement of IncRNAs in balancing the BMP signaling responses just started being explored. A recent study emphasized the regulation of BMP signaling by the IncRNA *PVT1* in glioma. The mechanism of action of *PVT1* involves competitive binding to *miR-128-3p*, which prevents it from binding to the *GREM1* mRNA, thereby stabilizing GREM1 levels, which antagonize the BMP2 and BMP4 ligands, leading to attenuation of BMP signaling (Fu et al., 2018a).

1.7 BMP-mediated physiological responses

BMP signaling participates in diverse physiological processes, such as differentiation, development and bone homeostasis (Davis et al., 2016). Initial studies of BMPs demonstrated a crucial role during bone formation (Sánchez-Duffhues et al., 2015). Many BMPs play roles during early development, defining the front-back and left-right polarity of the embryo (De Robertis, 2009). A well-known physiological function of BMP signaling is the induction of osteoblastic differentiation of the mouse myoblast cell line C2C12 (Katagiri et al., 1994). Moreover, BMP signaling is important for maintaining proper homeostasis of the cardiovascular system and the liver (Gomez-Puerto et al., 2018). Similar to TGFβ, the BMPs can also arrest the cell cycle and induce apoptosis via mechanisms that resemble those described above for TGFβ (Kawamura et al., 2002).

In glioblastoma multiforme (GBM), BMPs promote cancer stem cell differentiation towards astrocyte-like tumor cells, and reduce cell proliferation, inhibiting GBM progression (Caja et al., 2015). The BMP signaling also exerts important functions during the reversion of EMT and the establishment of mesenchymal-to-epithelial transition (MET). BMP induces the expression of Id proteins, which, in turn, suppress the expression of EMT-related transcription factors, shifting the balance from EMT to MET (Tan et al., 2015). How-
ever, reports also indicate that BMPs can induce the EMT under certain cancer-related contexts. For example, the BMP2 ligand promotes EMT of breast cancer cells, through PI3K/Akt-mediated degradation of the retinoblastoma (Rb) protein and induction of CD44 expression (Huang et al., 2017). Similarly to its effects in breast cancer, BMP2 exerts pro-EMT functions in gastric cancer cells. Activation of the BMP pathway results in induction of EMT markers and a concomitant repression of E-cadherin expression, an effect which is abolished by blockade of PI3K/Akt pathway (Kang et al., 2010). In addition, BMP signaling potentiates invasion of triple negative breast cancer cells, \textit{in vitro}, and bone metastasis in xenograft models (Katsuno et al., 2008).
2. Long non-coding RNAs

2.1 Non-coding RNAs

RNAs are versatile molecules which play crucial roles in basic cellular processes, such as transcription, splicing, translation and ribosome biogenesis. The role of messenger RNAs (mRNAs) as mediators of genetic information from DNA to proteins is highlighted at the central dogma of biology. According to this hypothesis, mRNAs transfer the genetic information which is maintained in DNA to polypeptides, which are the effector molecules of multiple cellular processes. The protein-coding mRNAs account for only a minor fraction of the genome, i.e., approximately 2%. However, the vast majority of the human genome consists of non-coding regions, such as pseudogenes, transposons, repeated sequences, and a large proportion is transcribed to non-coding RNAs (ncRNAs) with largely unknown functions. Among different types of non-coding RNAs, long non-coding RNAs (lncRNAs), which are arbitrarily defined as being longer than 200 nucleotides, represent the largest class of non-coding RNAs. This definition distinguishes lncRNAs from short non-coding RNAs, such as microRNAs (miRNAs), small interfering RNAs (siRNAs) and PIWI-interacting RNAs (piRNAs), whose roles in regulating gene expression are well characterized (Esteller, 2011). Recently another class of non-coding RNAs, which derive from back-splicing of mRNAs, has been described. These so called circular RNAs (circRNAs) have circular structure and often interact with miRNAs to modulate gene expression (Li et al., 2018). Long non-coding RNAs show both nuclear and cytoplasmic localization and their molecular functions depend on their subcellular distribution (Tordonato et al., 2015). Interestingly, lncRNAs, similarly to miRNAs and proteins can also be packed to exosomes and secreted in neighboring cells, where they exert their biological functions (Sun et al., 2018). Long non-coding RNAs can be capped, spliced and poly-adenylated similar to mRNAs, but they lack an open reading frame (ORF) or they contain only small ORFs which are not capable of producing a functional polypeptide (Hu et al., 2012).
2.2 Classification of IncRNAs

Although cases of IncRNAs in the context of TGFβ family signaling have already been presented, various structural and functional details of this emerging group of RNAs are discussed in this section. The class of IncRNAs can often be categorized according to their genomic location with respect to neighboring protein-coding genes. Thus, IncRNAs are divided in a) antisense, b) intronic, c) divergent, and d) intergenic subgroups (Rinn and Chang, 2012). Antisense IncRNAs are those transcribed from the opposite strand to that of the sense protein coding gene; they sometimes overlap with either the 5´- or the 3´- region of the sense transcripts (Pelechano and Steinmetz, 2013). Antisense IncRNAs can be further subdivided in different groups according to their proximity between the sense protein coding gene and the antisense transcript. Thus, these subgroups consist of: i) nearby to head transcripts, when the 5´ end of the sense mRNA is located close to the 5´ end of the antisense RNA, ii) nearby to tail, when the 3´ regions of both RNAs are nearby, iii) head to head, when the 5´ ends of sense and antisense RNAs show partial complementarity with each other, iv) tail to tail, which is similar to subgroup iii but refers to the 3´ ends of both RNAs, and v) full overlapping when there is a total overlap between the two transcripts (Villegas and Zaphiropoulos, 2015). The intronic IncRNAs arise from the intronic sequences of protein coding genes and do not overlap with exons. Divergent IncRNAs share the same promoter with a protein coding gene and are transcribed in the opposite direction to that of the coding gene but they do not overlap with it. Intergenic IncRNAs are located in distant regions from coding genes and they possess their own transcriptional units (Rinn and Chang, 2012). Another type of IncRNAs, termed enhancer RNAs (eRNAs), has also been described. These transcripts arise from enhancer elements and frequently regulate the expression of nearby genes in a positive way (Chen et al., 2017a). This group of IncRNAs, which is also designated as ncRNA-activating (ncRNA-a), directly interacts with the multiprotein complex Mediator, an important component for gene transcription, to activate transcription of neighboring genes, via chromosomal looping (Lai et al., 2013).

2.3 Biological functions of IncRNAs

Since their discovery, different cellular functions have been attributed to IncRNAs. They are involved in processes, such as genomic imprinting, X chromosome inactivation, self-renewal capacity, differentiation, development, as well as in pathological conditions, like tumor development and metastasis. Genomic imprinting is the process whereby a gene is expressed by one chromosome depending on the parental origin. Several proteins and
lncRNAs, such as Igf2r/Air and Kcnq1/Kcnq1ot1 are transcribed from imprinted loci and these lncRNAs epigenetically control the expression of nearby protein-coding genes (Kung et al., 2013).

Another process in which lncRNAs play important roles is the X chromosome inactivation (XCI), a phenomenon whereby one of the two X chromosomes is silenced in female mammals, in order to compensate for the genetic material between males and females. During XCI, one X chromosome in females is epigenetically silenced due to the function of the lncRNA Xist, which covers the whole X chromosome and recruits polycomb repressor complexes. The expression of Xist is regulated by neighboring lncRNAs, such as Tsix and Jpx; its recruitment to the X chromosome is mediated by the Yin Yang 1 (YY1) protein, which binds both DNA and Xist lncRNA, via distinct domains in YY1 (Jeon and Lee, 2011).

Additionally, lncRNAs are implicated in cell differentiation. For example, the lncRNA definitive endoderm-associated lncRNA 1 (DEANR1) induces expression of the transcription factor FOXA2, by assisting SMAD2/3 binding to the FOXA2 promoter, thus promoting human endoderm differentiation (Jiang et al., 2015). Moreover, Wang et al. showed that lncTCF7 is a tumor-promoting lncRNA, which is important for self-renewal of human liver cancer stem cells. Its mode of action is to control the expression of the transcription factor TCF7, which ultimately results in Wnt signaling activation and enhances cancer stem cell self-renewal capacity (Wang et al., 2015).

The contribution of lncRNAs in the regulation of immune responses has been appreciated during the last years. Several lncRNAs participate in the differentiation of hematopoietic stem cells and myeloid progenitor cells to generate specific immune cells. Inflammatory responses can also be modulated by lncRNAs, in a positive or negative way. Examples of pro-inflammatory lncRNAs are the lncRNA-Cox2 and PACER (Chen et al., 2017b). Lnc-Cox2 is up-regulated in response to Toll-like receptor (TLR) activation and regulates the expression of hundreds of inflammatory-related genes, in cooperation with the heterogeneous nuclear ribonucleoproteins hnRNP-A/B and hnRNP-A2/B1 (Carpenter et al., 2013). The expression of the lncRNA PACER is enhanced by lipopolysaccharide treatment and modulates the gene encoding the inflammatory regulator protein COX-2. PACER serves as a decoy for the NF-κB subunit p50 and facilitates the transition from repressive p50-p50 homodimer complexes to active p50-p65 NF-κB heterodimers, which positively regulate the gene encoding COX-2 (Krawczyk and Emerson, 2014). In contrast to the previous examples, lncRNAs can also limit inflammatory responses. For example, the IL-1β- and TNF-induced lncRNA Lethe exhibits negative effects on NF-κB signaling, by acting as a decoy for p65, thereby preventing p65 from binding to its target genes. This negative feedback loop restricts NF-κB-induced inflammation (Rapicavoli et al., 2013).
Another biological function in which lncRNAs participate is the programmed cell death or apoptosis. One of the most important effectors for transcriptional induction of the pro-apoptotic program is the transcription factor p53. The tumor suppressor protein p53 is activated in response to DNA damage and orchestrates defensive responses, exemplified by cell cycle arrest and apoptosis. Interestingly, p53 transcriptionally induces the lincRNA-p21, in order to execute its repressive functions on anti-apoptotic genes. LincRNA-p21 is an important effector of p53 for the establishment of apoptosis, although it does not affect p53-mediated cell cycle arrest. Mechanistically, lincRNA-p21 promotes repression of target genes, through physical interaction with the heterogeneous nuclear ribonucleoprotein K (hnRNP-K), which leads to the formation of a repressive RNA-protein complex at the promoters of target genes (Huarte et al., 2010).

2.4 Molecular functions of lncRNAs

Although the mechanisms of action of lncRNAs require further investigation, recent studies have attempted to unravel their functions. The molecular mechanisms by which lncRNAs perform their functions are largely determined by their subcellular distribution (Tordonato et al., 2015). The nuclear lncRNAs are usually involved in regulating gene expression at a transcriptional or post-transcriptional level. In addition, many nuclear lncRNAs participate in chromatin remodeling, through direct interaction with promoter or enhancer regions or by interacting with chromatin-modifying enzymes, such as histone methyltransferases, histone acetyltransferases or demethylases or DNA methyltransferases (Marchese et al., 2017). The cytoplasmic lncRNAs mainly function through interactions with cytoplasmic proteins, mRNAs or miRNAs.

2.4.1 Nuclear lncRNAs can function in cis or in trans

Nuclear lncRNAs, such as natural antisense lncRNAs, can act either in cis or in trans to regulate gene expression. When acting in cis, lncRNAs exert their functions at a restricted region nearby the site of their transcription, in other words they regulate the expression of genes in close proximity to the newly synthesized lncRNA. On the other hand, trans-acting lncRNAs are capable of acting at distant loci, for example other chromosomes, thus affecting the expression of genes far away from their transcription start site (Lee, 2012). The cis-acting lncRNAs can exert their functions in different ways. First, lncRNAs can modulate gene expression locally, by recruiting transcription factors or modifying the chromatin conformation on promoters of neighboring genes. Second, the act of transcription of a lncRNA can influence the transcription of a neighboring gene. In this case the sequence of the lncRNA is not important.
for the regulation of nearby genes. The third possibility is that the cis-regulatory function of a lncRNA depends on DNA sequences, which reside in its locus and does not involve the lncRNA transcript itself (Kopp and Mendell, 2018). The trans-acting lncRNAs perform their molecular roles in at least three ways. First, they are able to modulate expression of distal genes by altering the epigenetic landscape or the transcriptional machinery of target promoters. Second, they affect nuclear organization, by residing at specific nuclear subregions, due to associations with proteins, thereby creating a favorable microenvironment for progression of specific molecular functions, such as splicing and transcription. Finally, they bind to proteins or other types of RNA and influence the activity or the interaction of their molecular partners with other molecules (Kopp and Mendell, 2018).

2.4.2 LncRNAs are involved in transcriptional and epigenetic regulation of gene expression

In general, regulatory roles in transcription, chromatin remodeling, mRNA splicing, mRNA stability and translation have been attributed to many lncRNAs. According to a pioneering review by Wang and Chang, nuclear lncRNAs can be grouped as signal, decoy, guide, or scaffold molecules, depending on their mode of action. Several lncRNAs are expressed in a tissue- or time-dependent manner or in response to developmental or environmental cues, thus acting as signaling molecules (Wang and Chang, 2011). Moreover, lncRNAs act as scaffold molecules by interacting with two or more different proteins and bringing them to close proximity, thereby facilitating the formation of ribonucleoprotein complexes (Figure 4A). Furthermore, they guide proteins, such as transcription factors or components of chromatin remodeling complexes, to promoter sequences, providing specificity for these proteins to recognize their genomic targets (Figure 4B). Additionally, they are capable of titrating transcription factors or chromatin modifiers away from their target sequences in the genome, acting as decoys for these proteins (Figure 4C) (Wang and Chang, 2011). Recently, novel molecular mechanisms have been attributed to lncRNAs. For example, lncRNAs can directly bind to DNA strands, at sites of transcription, thereby creating DNA-RNA triplex formations and making specific chromatin regions accessible to transcription factors (Figure 4D). In a different mechanism, lncRNAs can positively influence gene transcription by bridging enhancer and promoter regions, which normally are located to distant genomic areas. The concomitant function of enhancer-derived lncRNAs, transcription factors and chromosomal looping facilitates the transmission of signals from enhancers to the proximal promoter regions for transcription initiation (Figure 4E).

As mentioned before, lncRNAs are implicated in transcriptional regulation. This can be achieved in many ways, one of the best characterized being the
epigenetic regulation via modulation of the chromatin landscape. There are several examples of lncRNAs interacting with histone-modifying complexes, such as the HOX transcript antisense RNA (HOTAIR), the antisense ncRNA in the INK4 locus (ANRIL) and the KCNQ1 opposite strand or antisense transcript 1 (KCNQ1OT1). The lncRNA-protein interaction assists histone-remodeling complexes in their recruitment to specific genomic loci (Geisler and Coller, 2013).

2.4.3 LncRNAs mediate post-transcriptional regulation

The role of lncRNAs during alternative splicing is of high importance. The expression profiles of alternatively spliced RNAs can frequently be cell-type specific, as exemplified by splicing of the fibroblast growth factor receptor 2 (FGFR2) gene. In this example, alternative splicing is coupled to the histone methylation landscape of the gene, which determines the recruitment of splicing factors and the subsequent inclusion or exclusion of the exon IIIb from the synthesized mRNA (Gonzalez et al., 2015). In human mesenchymal stem cells, the FGFR2 gene shows high di- or trimethylation of lysine 36 at histone H3 (H3K36me2,3), which prevents inclusion of exon IIIb. On the other hand, in epithelial cells H3K27me3 is abundant, which favors the inclusion of exon IIIb. Interestingly, an antisense RNA from the FGFR2 (asFGFR2) locus is generated in the presence of exon IIIb, which mediates the recruitment of Polycomb repressor complex proteins and the histone demethylase KDM2a, in order to block binding of the negative splicing factor PTB, which splices exon IIIb. Thus, asFGFR2 facilitates an epithelial-specific alternative splicing of FGFR2, by acting as a scaffold molecule for the association of Polycomb repressor proteins and KDM2a to the FGFR2 locus (Gonzalez et al., 2015).

Moreover, lncRNAs can affect gene expression at the post-transcriptional level. One such example is the lncRNA antisense Uchl1, which is a 5’ head to head antisense RNA to the gene encoding the mouse ubiquitin carboxy-terminal hydrolase L1 (Uchl1) (Carrieri et al., 2012). Antisense Uchl1 enhances the protein levels of Uchl1 without affecting its mRNA transcripts. The function of this lncRNA is mediated by a SINEB2 element, which is important for the recruitment of Uchl1 mRNA in polysomes for translation. Furthermore, the 5’ sequence overlapping with the sense Uchl1 mRNA sequence is another important feature allowing the interaction between the sense mRNA and the antisense lncRNA, due to complementarities in their sequences, thereby facilitating targeting of Uchl1 mRNA by its antisense partner (Carrieri et al., 2012).

Although regulation by lncRNAs at the pre- and post-transcriptional levels is better understood, there is less evidence for a potential role of lncRNAs during post-translational modifications of proteins. However, a recent study described a role of the lncRNA HOTAIR in the ubiquitination of Ataxin-1 and Snurportin-1 proteins, resulting in their proteasome-mediated degradation. Specifically, HOTAIR was shown to act as a mediator molecule that brings
Ataxin-1 and Snurportin-1 together with their corresponding E3 ubiquitin ligases and promotes their destruction, in order to block pre-mature senescence (Yoon et al., 2013).

2.4.4 Cytoplasmic IncRNAs

The IncRNAs that are predominantly localized at the cytoplasm can modulate gene expression and signaling pathways through diverse mechanisms. A widespread molecular function of IncRNAs at the cytoplasm derives from the “competing endogenous RNA” hypothesis (Salmena et al., 2011). According to this, IncRNAs bind to specific miRNAs and sequester away miRNAs from their mRNA targets. The specificity for the IncRNA-miRNA interaction is provided by sequence complementarity between IncRNA and miRNA. This mechanism has gained a lot of attention and includes many examples of IncRNA-miRNA partners (Figure 4F). Cytoplasmic IncRNAs can also facilitate protein translocation into the nucleus, through interaction with specific proteins of the nuclear pores, thereby modulating signaling responses (Figure 4G). The stability of mRNAs is another mechanism, by which IncRNAs regulate gene expression. IncRNAs can interact with specific mRNA regions, such as 5’- or 3’-UTRs and affect mRNA stability (Figure 4H). Signaling pathways can be modulated by cytoplasmic IncRNAs, which bind to effector proteins and interfere with their post-translational modifications, such as phosphorylation, that ultimately lead to the activation of pathway-dependent responses (Figure 4I).

2.5 Polycomb repressor complex 2 (PRC2) and IncRNAs interplay

Although every cell type shares the same copies of genes, gene expression profiles are tightly regulated, so that particular cell types express only the genes that enable them to perform their specific biological function. Moreover, gene expression regulation provides distinct expression patterns, required for the formation of different tissues, during embryonic development. Chromatin remodeling plays major roles in controlling cellular transcriptional programs. Open chromatin conformations are accessible to transcription factors and are generally considered as transcriptionally active genomic loci. In contrast, “closed” chromatin, such as heterochromatin, is characterized by high condensation and absence or low levels of transcriptional activity.

Epigenetic silencing of gene expression is achieved e.g. by DNA methylation, histone methylation, histone deacetylation or histone ubiquitination. Histone methylation can positively or negatively contribute to the regulation of
transcription; this depends on the specific amino acid residues that are modified. Examples of activating histone modifications include the trimethylation of lysine 4 at histone H3 (H3K4me$^3$) and trimethylation of lysine 36 at histone H3 (H3K36me$^3$). On the other hand, repressive chromatin is enriched in H3K9me$^3$ and H3K27me$^3$.

One of the protein complexes identified as mediators of epigenetic silencing is the polycomb repressive complex 2 (PRC2) (Margueron and Reinberg, 2011). The PRC2 complex mediates epigenetic silencing, through establishing the H3K27me$^3$ marker at promoter regions, thereby promoting compact chromatin structure. The core members of the PRC2 complex are the methyltransferase enhancer of zeste 2 (EZH2) and the adaptor proteins SUZ12, embryonic ectoderm development (EED) and RbAp46/48. These four proteins are required for the catalytic activity of PRC2 in vitro. Moreover, additional proteins such as JARID2, AEBP2 and Polycomb-like proteins (PCLs) are facultative subunits of the PRC2 complex and modulate its catalytic activity or its recruitment across genome. In fact, the interaction of the four core subunits with different combinations of the facultative subunits gives rise to two forms of PRC2 complex, the PRC2.1 and PRC2.2. The PRC2.1 complex includes one of the PCL proteins (PHF1, MTF2 or PHF19) and either the protein EPOP or the protein C10orf12. The PRC2.2 complex consists of the four core components together with the subunits AEBP2 and JARID2 (Holoch and Margueron, 2017). The PRC2 complexes promote silencing of a wide range of genes; however, a substantial amount of work suggests an important role of PRC2 in the repression of developmental genes, which maintains pluripotency. In addition, PRC2 positively contributes to cell proliferation by silencing the expression of cell cycle inhibitors, such as p15 and p16. Thus, members of the PRC2 complex, such as the catalytic subunit EZH2 are often overexpressed in various cancers (Margueron and Reinberg, 2011).

The recruitment of PRC2 to specific genomic loci is still debatable. PRC2 usually binds GC-rich genomic regions, such as CpG islands at transcriptionally inactive loci. Since specific DNA-binding motifs for PRC2 complex have not been identified, various hypotheses have been suggested, in order to explain the association of PRC2 with specific genomic regions. One of them, emphasizes the function of a closely related complex to PRC2, designated as PRC1, as a facilitator of PRC2’s recruitment. PRC1 establishes epigenetic silencing by monoubiquitinating H2AK119. This modification may assist PRC2’s recruitment to the genome, in a JARID2-dependent manner. According to this model, PRC2 can expand the epigenetic silencing, initiated by PRC1, by depositing H3K27me$^3$ modifications, which are recognized by the adaptor protein EED, thereby enhancing the activity of the complex and propagating repressive chromatin (Holoch and Margueron, 2017). Alternatively, lncRNAs can mediate specific recruitment of the PRC2 complex to genomic regions. In fact, several lncRNAs have been shown to interact with the subu-
nits EZH2 and SUZ12 (Khalil et al., 2009). In addition, the epigenetic silencing observed in the cases of the X chromosome inactivation and the repression of the developmental genes of the HOXD locus, is attributed to the activity of PRC2 complexes that are directed to their target genomic regions by the lncRNAs Xist and HOTAIR, respectively. Many more examples of lncRNA-PRC2 associations support this model in the literature. On the other hand, the increasing number of PRC2-associated lncRNAs and the lack of specific RNA-binding motifs on EZH2 or SUZ12 protein sequences raise the question about the level of specificity between lncRNAs and PRC2 interactions (Davidovich and Cech, 2015). Nevertheless, the interaction of nuclear lncRNAs with PRC2 complexes and the ability of lncRNAs, with GC-rich sequences, to directly bind to genomic loci, through DNA-RNA triplex formation, makes lncRNAs attractive candidates for directing PRC2 to the correct genomic regions.

Figure 4. Molecular functions of lncRNAs. Nuclear lncRNAs act as scaffold (A), guide (B) or decoy (C) molecules. They also form triple helix formations by directly binding to DNA (D), and act as enhancer-like RNAs, promoting gene transcription (E). Cytoplasmic lncRNAs function as sponges for miRNAs (F), facilitate nuclear-cytoplasmic protein translocation (G), affect mRNA stability (H) and interfere with post-translational modifications of proteins (I). TF: transcription factor, MED: component of the mediator complex, RNA pol II: RNA polymerase II, CDS: coding sequence.
2.6 TGFβ-regulated lncRNAs

The involvement of lncRNAs in TGFβ signaling has been studied over the last few years. The number of lncRNAs that are targeted by TGFβ stimulation is increasing rapidly. For examples see Table 1.

Richards et al. (2015) identified lncRNA-HIT (HOXA transcript induced by TGFβ) as a TGFβ target gene in NMuMG mouse mammary epithelial cells. This lncRNA belongs to the HOXA gene cluster and mediates TGFβ-induced EMT by targeting E-cadherin as its depletion impairs migration and invasion of NMuMG cells and rescues E-cadherin mRNA and protein levels. In addition, lncRNA-HIT is up-regulated in highly metastatic mouse and invasive human tumors.

In another study, Yuan et al., described a lncRNA-activated by TGFβ (lncRNA-ATB) which is induced by TGFβ stimulation of hepatocellular carcinoma cells (HCC), for several days (Yuan et al., 2014). LncRNA-ATB induces EMT, invasion and metastasis in HCC. Mechanistically, it increases ZEB1 and ZEB2 levels, due to its role as a ceRNA for miR-200s, which normally target ZEB1 and ZEB2 mRNAs for degradation. Moreover, lncRNA-ATB promotes metastatic organ colonization by interacting with interleukin-11 (IL-11) mRNA, thereby enhancing IL-11 expression and STAT3 signaling which is an important feature for survival of metastatic cells.

During TGFβ-mediated EMT, the lncRNA lnc-Spry1 is down-regulated in response to TGFβ treatment. Lnc-Spry1 is a negative regulator of EMT, as its depletion leads to elevated migration and invasion. At the molecular level, lnc-Spry1 interacts with the splicing factor U2AF65 and affects the alternative splicing of fibroblast growth factor receptors (Rodríguez-Mateo et al., 2017).

In a recent report, the liver fibrosis-associated lncRNA1 (lnc-LFAR1), which is specifically expressed in liver cells, was found to be a promoter of liver fibrosis, a function that is achieved through the activation of TGFβ and Notch signaling pathways (Zhang et al., 2017). The expression of lnc-LFAR1 is induced by TGFβ stimulation and in turn, this lncRNA potentiates the interaction between the TGFβR1 and Smad2/3, resulting in enhanced phosphorylation of Smad2/3. In addition, lnc-LFAR1 binds Smad2/3 and enhances the expression of TGFβ and Notch related genes (Zhang et al., 2017).

Profiling of TGFβ-target genes in intrahepatic cholangiocarcinoma (iCCA) revealed a novel TGFβ-induced long noncoding RNA (TLINC). TLINC encodes a long and a short isoform, which are found both in the nucleus and in the cytoplasm. Expression of the long TLINC isoform correlated with migratory behavior and with induction of proinflammatory cytokines in iCCA cells (Merdrignac et al., 2018).
<table>
<thead>
<tr>
<th>LncRNA</th>
<th>Type of regulation</th>
<th>Cell line/Tissue</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>IncRNA-ATB</td>
<td>Induction</td>
<td>HCC</td>
<td>(Yuan et al., 2014)</td>
</tr>
<tr>
<td>IncRNA-HIT</td>
<td>Induction</td>
<td>NMuMG</td>
<td>(Richards et al., 2015)</td>
</tr>
<tr>
<td>MALAT1 has2as</td>
<td>Induction</td>
<td>Bladder cancer</td>
<td>(Fan et al., 2014)</td>
</tr>
<tr>
<td>Inc-LFAR1</td>
<td>Induction</td>
<td>Liver</td>
<td>(Zhang et al., 2017)</td>
</tr>
<tr>
<td>Inc-Spry1</td>
<td>Reduction</td>
<td>NMuMG</td>
<td>(Rodriguez-Mateo et al., 2017)</td>
</tr>
<tr>
<td>TLINc</td>
<td>Induction</td>
<td>Cholangiocarcinoma</td>
<td>(Merdrignac et al., 2018)</td>
</tr>
<tr>
<td>MIR100HG</td>
<td>Induction</td>
<td>PDAC</td>
<td>(Ottaviani et al., 2018)</td>
</tr>
<tr>
<td>ANCR</td>
<td>Reduction</td>
<td>MCF10A</td>
<td>(Li et al., 2017)</td>
</tr>
<tr>
<td>LINC01186</td>
<td>Reduction</td>
<td>A549</td>
<td>(Hao et al., 2017)</td>
</tr>
<tr>
<td>H19</td>
<td>Reduction</td>
<td>Tumor-initiating hepatocytes</td>
<td>(Zhang et al., 2018b)</td>
</tr>
<tr>
<td>H19</td>
<td>Induction</td>
<td>Hep3B</td>
<td>(Matouk et al., 2014)</td>
</tr>
<tr>
<td>TBIIA</td>
<td>Induction</td>
<td>NSCLC</td>
<td>(Lu et al., 2018a)</td>
</tr>
<tr>
<td>EPB41L4A-AS2</td>
<td>Reduction</td>
<td>HNSCC</td>
<td>(Huang et al., 2018)</td>
</tr>
<tr>
<td>Erbb4-IR</td>
<td>Induction</td>
<td>Renal fibrosis</td>
<td>(Feng et al., 2018)</td>
</tr>
<tr>
<td>IncRNA-LET</td>
<td>Reduction</td>
<td>Urinary bladder cancer</td>
<td>(Zhuang et al., 2017)</td>
</tr>
<tr>
<td>UCA1</td>
<td>Induction</td>
<td>Gastric cancer</td>
<td>(Zuo et al., 2017)</td>
</tr>
<tr>
<td>NKILA</td>
<td>Induction</td>
<td>Esophageal squamous cell carcinoma</td>
<td>(Lu et al., 2018b)</td>
</tr>
<tr>
<td>AC067945.2</td>
<td>Reduction</td>
<td>Normal skin fibroblasts</td>
<td>(Chen et al., 2018a)</td>
</tr>
<tr>
<td>RP6-65G23.5</td>
<td>Induction</td>
<td>MCF10A</td>
<td>(Liao et al., 2018a)</td>
</tr>
<tr>
<td>IncRNA-LINP1</td>
<td>Reduction</td>
<td>Lung cancer</td>
<td>(Zhang et al., 2018a)</td>
</tr>
<tr>
<td>MEG8</td>
<td>Induction</td>
<td>A549, LC2/ad cells, Panc1 cells</td>
<td>(Terashima et al., 2018)</td>
</tr>
<tr>
<td>Inc-TSI</td>
<td>Induction</td>
<td>Renal tubular epithelial cells</td>
<td>(Wang et al., 2018b)</td>
</tr>
</tbody>
</table>
2.7 LncRNAs as regulators of the TGFβ pathway

The low levels of evolutionary conservation that characterizes many lncRNAs, as well as their tissue-specific expression patterns make them useful context-dependent candidate regulators of TGFβ signaling. That said, it is difficult to consider lncRNAs as crucial determinants of the TGFβ signaling cascade as for example the members of the Smad family, which are conserved throughout all metazoan species. Nevertheless, lncRNAs could arise in the highly complex human genome to provide additional layers of regulation in a tissue- or disease-specific manner, in the already complicated regulatory circuits, that determine the outcome of the TGFβ pathway. Several lncRNAs modulate TGFβ signaling (Table 2); expression of some of these lncRNAs is also regulated by TGFβ signaling, thus forming positive or negative regulatory feedback loops.

One of the best characterized modulators of TGFβ signaling is the lncRNA maternally expressed gene 3 (MEG3), which is located at repressive chromatin regions together with EZH2, a component of the PRC2 repressor complex and is recruited to its target loci through RNA-DNA triplex formation (Mondal et al., 2015). MEG3 together with EZH2 regulate common target genes, including TGFβ-related genes by binding to distal regulatory elements in their promoters. Furthermore, siRNA-mediated knock-down of MEG3 results in enhanced expression of SMAD2, TGFB2 and TGFBR1, whereas ectopic expression of MEG3 decreases the expression levels of these genes, suggesting that MEG3 is a negative regulator of TGFβ signaling (Mondal et al., 2015).

In addition, the lncRNA growth arrest-specific 5 (GAS5) inhibits TGFβ signaling during smooth muscle differentiation. GAS5 exerts its inhibitory function through interaction with Smad3, which prevents the binding of Smad3 to its genomic loci. GAS5 can bind Smad3 via rSBEs, which reside in the sequence of this lncRNA, thus acting as a decoy molecule for Smad3 (Tang et al., 2017).

Another example of an lncRNA with a regulatory role in TGFβ-induced smooth muscle differentiation is the brain cytoplasmic RNA 1 (BC1). The canonical TGFβ signaling is disrupted by BC1 over-expression, as BC1 interacts with rSBEs of Smad3 and retains it at the cytoplasm. The prevention of Smad3 translocation into the nucleus by BC1 leads to decreased expression of smooth muscle cell differentiation proteins, such as calponin and smooth muscle α-actin (α-SMA) (Wang et al., 2018c).

The lncRNA NORAD is an example of an lncRNA that positively regulates TGFβ signaling. NORAD is located at the cytoplasm and promotes the transport of Smad3 into the nucleus, through facilitating the interaction of importin β1 with Smad3 in A549 lung adenocarcinoma cells. Thus, NORAD also potentiates EMT, by enhancing TGFβ signaling (Kawasaki et al., 2018).

The importance of lncRNAs in the TGFβ-mediated T regulatory cells polarization is exemplified by the finding of the lnc-Smad3, an lncRNA that
binds to the histone deacetylase HDAC1 and suppresses the transcription of SMAD3. The recruitment of HDAC1 at the promoter of SMAD3 prevents the binding of the histone methyltransferase Ash1l to the promoter of SMAD3, which positively regulates SMAD3 transcription. On the other hand, TGFβ signaling down-regulates lnc-Smad3 expression, making the promoter of SMAD3 accessible for Ash1l and thus neutralizes the negative regulation of SMAD3 transcription by lnc-Smad3 (Xia et al., 2017).

Table 2. List of lncRNAs that modulate TGFβ signaling

<table>
<thead>
<tr>
<th>LncRNA</th>
<th>Type of modulation</th>
<th>Cell line/Tissue</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>MEG3</td>
<td>Negative</td>
<td>BT-549 cells</td>
<td>(Mondal et al., 2015)</td>
</tr>
<tr>
<td>GAS5</td>
<td>Negative</td>
<td>C3H10T1/2 mesenchymal cells</td>
<td>(Tang et al., 2017)</td>
</tr>
<tr>
<td>BCI</td>
<td>Negative</td>
<td>Smooth muscle cells</td>
<td>(Wang et al., 2018c)</td>
</tr>
<tr>
<td>lnc-LFAR1</td>
<td>Positive</td>
<td>Liver</td>
<td>(Zhang et al., 2017)</td>
</tr>
<tr>
<td>NORAD</td>
<td>Positive</td>
<td>A549 cells</td>
<td>(Kawasaki et al., 2018)</td>
</tr>
<tr>
<td>LINK-A</td>
<td>Positive</td>
<td>Ovarian carcinoma</td>
<td>(Ma and Xue, 2018)</td>
</tr>
<tr>
<td>EPB41L4A-AS2</td>
<td>Negative</td>
<td>HNSCC</td>
<td>(Huang et al., 2018)</td>
</tr>
<tr>
<td>Erbb4-IR</td>
<td>Positive</td>
<td>Renal fibrosis</td>
<td>(Feng et al., 2018)</td>
</tr>
<tr>
<td>CCAT2</td>
<td>Positive</td>
<td>Breast cancer</td>
<td>(Wu et al., 2017)</td>
</tr>
<tr>
<td>BX357664</td>
<td>Negative</td>
<td>Renal cell carcinoma</td>
<td>(Liu et al., 2016)</td>
</tr>
<tr>
<td>ANRIL</td>
<td>Negative</td>
<td>Thyroid cancer</td>
<td>(Zhao et al., 2016)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Esophageal squamous cell carcinom</td>
<td>(Chen et al., 2014)</td>
</tr>
<tr>
<td>TUG1</td>
<td>Positive</td>
<td>Pancreatic cancer</td>
<td>(Qin and Zhao, 2017)</td>
</tr>
<tr>
<td>LINC00978</td>
<td>Positive</td>
<td>Gastric cancer</td>
<td>(Fu et al., 2018b)</td>
</tr>
<tr>
<td>SPRY4-IT</td>
<td>Positive</td>
<td>Thyroid cancer</td>
<td>(Zhou et al., 2018)</td>
</tr>
<tr>
<td>CHRF</td>
<td>Positive</td>
<td>Primary myocardial cells</td>
<td>(Chen et al., 2018b)</td>
</tr>
<tr>
<td>lnc-TSI</td>
<td>Negative</td>
<td>Renal tubular epithelial cells</td>
<td>(Wang et al., 2018b)</td>
</tr>
</tbody>
</table>
3. MicroRNAs

3.1 Biogenesis and functions of microRNAs

MicroRNAs (miRNAs) are short non-coding RNAs consisting of 22 nucleotides which are often generated by the intronic regions of protein-coding or non-coding genes (Ha and Kim, 2014). The production of miRNAs is a multistep process, which is achieved through the co-operative functions of transcription and RNA processing enzymes. During the first step, RNA polymerase II initiates the transcription of the host gene, in which an miRNA or a group of miRNAs reside resulting in the formation of a long primary miRNA (pri-miRNA) of more than 1 kb length. Then the pri-miRNA undergoes two rounds of cleavage by the ribonucleases Drosha and Dicer. The first cleavage of the pri-miRNA takes place in the nucleus and is catalyzed by the Microprocessor complex, which consists of the RNase III-type endonuclease Drosha and the cofactor DGCR8, an auxiliary protein that strengthens the binding of pri-miRNA to Drosha. The product of this first step of processing is a ~65 nucleotides hairpin RNA that is called precursor miRNA (pre-miRNA). The pre-miRNA is then exported to the cytoplasm, where it undergoes a second cleavage by the RNase III Dicer, generating the mature miRNA of approximately 22 nucleotides (Ha and Kim, 2014). After processing by Dicer, the mature miRNA is usually found in the form of duplexes, consisting of one passenger and one guide strand. The guide strand is captured by the miRNA-induced silencing complex (miRISC) and is directed to its mRNA target sequences, while the passenger strand is degraded (Ha and Kim, 2014).

Mature miRNAs target complementary 3′-UTR regions of mRNAs and mediate gene silencing at the post-transcriptional level. The recognition of mRNAs by miRNAs results in mRNA degradation or in interference of mRNA translation. The level of complementarity between miRNAs and mRNAs determines the mode of miRNA-mediated gene repression. Complete complementarity between a miRNA and its target mRNA favors mRNA cleavage. In contrast, partial complementarity of these two molecular partners shifts the balance towards inefficient mRNA translation (Bartel, 2004).

Since the recognition of mRNAs by miRNAs is redundant and the majority of the mammalian mRNAs possess conserved miRNA-binding sites, it is plausible to assume that miRNAs participate in the regulation of several biological processes (Bartel, 2009). Indeed, miRNAs play important roles in development, cell cycle progression, apoptosis, EMT and are often deregulated in
pathological conditions, such as cancer. Interestingly, there are examples of miRNA clusters, which consist of different miRNAs and are located in a unique genomic locus. The miRNA clusters are commonly regulated at the transcriptional level and often share common biological functions. Examples of miRNA clusters include the miR-17/92 cluster, which regulates mRNAs related to cell cycle progression and the miR-200 cluster, which antagonizes EMT (Hayes et al., 2014).

3.2 Micro-RNAs in TGFβ signaling

TGFβ signaling regulates the expression of several miRNAs, which mediate many aspects of its physiological roles. For example, TGFβ transcriptionally induces miR-100 and miR-125b expression, by enhancing the transcription of the primary MIR100HG lncRNA, in pancreatic ductal adenocarcinoma. Increased miR-100 and miR-125b levels promote a pro-EMT phenotype and enhance the tumor-initiating capacity of pancreatic cancer cells (Ottaviani et al., 2018). In addition, TGFβ up-regulates miR-99a/b, upon EMT induction in NMuMG cells (Chen et al., 2016). The expression of miR-181a is also enhanced in response to TGFβ, which leads to increased metastasis, through activating the PI3K/Akt pathway, increasing Erk phosphorylation and down-regulating E-cadherin (Taylor et al., 2013). Moreover, TGFβ increases the miR-491-5p levels leading to down-regulation of the polarity complex Par3 in rat proximal tubular epithelial cells (Zhou et al., 2010).

One of the best-studied examples of TGFβ-down-regulated miRNAs is the miR-200 family. The miR-200 family inhibits EMT, by targeting ZEB1 and ZEB2 mRNAs for degradation. A part of TGFβ’s program to establish EMT includes the down-regulation of miR-200, which leads to stabilization of ZEB1 and ZEB2 mRNAs (Gregory et al., 2011).

Several members of the TGFβ signaling pathway undergo post-transcriptional regulation by miRNAs. For example, members of the let-7 miRNA family target the TβRI and play important roles during development (Tzur et al., 2009). In addition, the miR-23b cluster down-regulates Smad3, Smad4 and Smad5 in fetal hepatocytes, while miR-21 targets Smad7, thereby enhancing TGFβ signaling (Butz et al., 2012). The miR-520/373 family reduces the levels of TβRII and attenuates invasion of breast cancer cells and their intravasation to the vasculature. Furthermore, the miR-155 down-regulates Smad2 and compromises TGFβ signaling in a monocytic line (Louafi et al., 2010). The miR17/92 cluster has an inhibitory role on TGFβ signaling due to the suppressive effects of miR-17-5p and miR-20 on TβRII and the down-regulation of Smad4 by miR-18 (Dews et al., 2010).
4. The poly(ADP)ribose polymerase family

4.1 Poly(ADP)ribosylation
Post-translational modification is the covalent attachment of chemical groups to specific amino acid residues in proteins after their translation. These modifications are important for controlling the activity of target proteins, regulating protein stability, signal transduction, transcriptional regulation, which regulate multiple cellular processes, such as cell proliferation, DNA damage response and apoptosis. The best characterized post-translational modifications are phosphorylation, acetylation, methylation, glycosylation, SUMOylation, ubiquitination and poly(ADP)ribosylation (PARylation) (Knorre et al., 2009). The latter is a widespread modification which is catalyzed by the enzymes of the poly(ADP)ribose polymerase (PARP) family. The enzymatic reaction requires nicotinamide adenine dinucleotide (NAD+) as a donor molecule for ADP-ribose units, which are transferred to Glu, Asp or Lys residues of acceptor proteins. This results in the attachment of poly(ADP)ribose polymers (PARs), consisting of a few to approximately 200 ADP-ribose units. The ADP-ribose residues are connected through 1→2 O-glycosidic bonds and form negatively charged, linear or branched polymers, thus affecting protein-protein or nucleic acid-protein interactions, protein localization, as well as providing a platform for recruitment of proteins in specific subcellular or genomic loci in response to different stimuli (Teloni and Altmeyer, 2016). The levels of PARylation are controlled by “writers” and “erasers” which modulate the deposition and degradation of the PAR polymers. The “writers” involve the PARPs, which catalyze the transfer of ADP-ribose monomers to acceptor proteins. The “erasers” consist of enzymes, such as the poly(ADP-ribose) glycohydrolase (PARG), which removes ADP-ribose units from PARylated proteins. PARylation is recognized by “readers”, referring to proteins with PAR-binding domains, which interact with PAR polymers and mediate PAR-dependent biological processes (Karlberg et al., 2013).

4.2 Poly(ADP)ribose polymerases (PARPs)
The PARP family consists of 17 members in humans, categorized into four subgroups, which are divided based on structural criteria (Gupte et al., 2017). The first group involves the exclusively nuclear and DNA-dependent PARPs
(PARP1, PARP2 and PARP3) which are activated by DNA breaks and interact with damaged DNA through their zinc finger domains. The second group, called tankyrases, mediates protein-protein interactions, whereas the third group of Cys-Cys-Cys-His (CCCH) PARPs binds RNA and proteins with WWE domains. The last group consists of the macroPARPs which recognize ADP-ribose units through their macrodomains, a property that enables them to be recruited at sites of PARylation (Gupte et al., 2017). Although the majority of PARylation takes place at the nucleus, 6 out of the 17 human PARPs (PARP5a, PARP7, PARP10, PARP12, PARP13 and PARP15) are located at the cytoplasm, which may indicate that PARylation affects molecular processes that are executed at the cytoplasm (Bai, 2015). All members of the PARP family share an evolutionary conserved catalytic domain, in which a highly conserved PARP signature motif is located. This motif forms the active site, where NAD$^+$ molecules bind to the “donor site”; perturbations in the conserved His-Tyr-Glu tripeptide lead to abrogation of the enzymatic activity of PARPs. In addition to the donor motif, another conserved sequence, which is called “acceptor site”, is necessary for binding of acceptor proteins (Barkauskaite et al., 2015). Under stress conditions, PARP1 catalyzes the vast majority of cellular PAR polymers, suggesting that it is an important enzyme for PARylation. Although PARP2 is characterized by lower catalytic activity compared to PARP1, it partially compensates for PARP1 activity in PARP1 knock-out mouse embryonic fibroblasts (Yelamos et al., 2011). While PARP1 and PARP2 represent the enzymes with the highest PARylation activity, other members of the family also contribute to molecular and biological processes. For example, PARP3 and PARP14 participate in the modulation of genes related to inflammatory responses. Moreover, the cytoplasmic PARPs are involved in the regulation of the mRNA fate at the cytoplasm (Bai, 2015). Nevertheless, although there is a substantial understanding of the molecular functions of PARP1 and PARP2, the function of other members of the PARP family (i.e. PARP8, PARP11) is still not yet elucidated.

4.3 Catabolism of Poly(ADP) ribosylation

The catalysis of PARylation by PARPs is a dynamic process which is tightly regulated so that immediate responses to intra- or extra-cellular signals are quenched as soon as the outcome has been attained. For example, in response to DNA damage caused by genotoxic stress, PARPs PARylate several members of the DNA repair machinery, in order to be recruited to the sites of DNA double strand breaks and repair the damage (Li and Yu, 2015). After reparation is completed, the PAR polymers are rapidly degraded; otherwise the accumulation of PAR polymers triggers cell death signals, due to NAD$^+$ and ATP depletion. The loss of NAD$^+$ and ATP pools is a consequence of PARP1 hyperactivation, which ultimately leads to necrotic cell death, characterized
by cell swelling and rupture. Thus, the catabolism of PARs is of high importance for normal cell homeostasis. The hydrolysis of PARs is performed by the enzyme PARG which possesses both endo- and exo-glycohydrolase activities (Mashimo and Moss, 2016). This enzyme is responsible for the glycohydrolysis of the $1\rightarrow2$ glycosidic bond between the ADP-ribose residues, thus producing free mono(ADP)ribose or oligo(ADP)ribose units. The requirement of PARG for normal cell physiology and survival is highlighted by the observation that PARG null mice die at an early embryonic stage (Koh et al., 2004).

A single PARG gene encodes different splice variants which are translated to several isoforms, with different subcellular localizations. The two most abundant isoforms are a long (110 kDa) nuclear/cytoplasmic isoform and a 65 kDa cytoplasmic isoform, both being catalytically active enzymes. Additionally, there are 102 kDa and 99 kDa isoforms which are cytoplasmic, and a 55 kDa isoform which is found in mitochondria (Niere et al., 2012). The existence of cytoplasmic PARG isoforms may indicate that de(ADP)ribosylation is also performed in the cytoplasm, although PARPs are mainly nuclear enzymes. Other PAR-degrading enzymes are the proteins MacroD1 and MacroD2, the members of the ADP-ribosyl-acceptor hydrolases (ARHs) and the terminal ADP-ribose protein glycohydrolase 1 (TARG1) (Bu et al., 2018). The ADP-ribosyl hydrolase 3 (ARH3) which degrades PAR chains mainly in mitochondria and, in contrast to PARG, catabolizes PAR rather inefficiently in vitro (Niere et al., 2012).

4.4 Cellular functions of PARPs

The molecular structure of PAR polymers and the dynamic process of their polymerization and de-polymerization offer a mechanism for immediate and transient control of diverse cellular processes. Indeed, PARPs’ functions are crucial for DNA damage response, transcription regulation, chromatin dynamics and cell death pathways (Gupte et al., 2017). The role of PARP1 in DNA repair is well characterized. PARP1 recognizes single or double strand breaks, binds to them through its zinc-finger domain and interacts with several components of the DNA repair machinery, such as XRCC-1 and DNA-dependent protein kinase. PARP1 auto-PARylation can also function as a protein recruitment platform, facilitating the attachment of DNA repair proteins at the sites of DNA damage. Thus, PARP1 participates in base excision repair (BER), nucleotide excision repair (NER) and double strand break (DSB) repair pathways (Beck et al., 2014). The other DNA-dependent PARP (PARP-2) shows redundant function with PARP1 and is also involved in DNA repair mechanisms, although the fact that PARP-2 is recruited to damaged DNA with slower kinetics and more persistently compared to PARP1, suggests that it
plays its most important role at later stages of the repair processes (Yelamos et al., 2011).

Although PARPs can act as guardians of genome maintenance under low levels of DNA damage, they are also involved in cell death processes during excessive genomic damage. During severe breakage of DNA strands, cells are incapable of repairing the damage and they ultimately die either in an orchestrated manner via apoptosis, or through necrosis, whereby cells swell due to ATP depletion and inflammatory responses are induced in the surrounding microenvironment. Interestingly, JNK activation and calpain-mediated activation of Bax are important steps in induction of necrosis (Douglas and Baines, 2014). Alternatively, PARP1 hyper-activation leads to release of apoptosis-inducing factor (AIF) from mitochondria and its translocation to the nucleus, where it promotes DNA fragmentation and chromatin condensation. On the other hand, AIF release from mitochondria is also induced by the fatty acid α-eleostearic acid (α-EA) in a PARP1-independent fashion; MEK kinase is an important player in this process (Kondo et al., 2010). In summary, PARPs present crucial roles in regulating DNA repair, cell death and signaling pathway mechanisms (Figure 5A). The biological roles of PARPs are mediated by their ability to ADP-ribosylate target proteins, which allow them to recruit proteins at specific genomic regions, affecting chromatin remodeling, to regulate transcription, to alter protein localization or interfere with nucleic acid-protein interactions (Figure 5B).

4.5 PARPs in TGFβ signaling

The involvement of the PARP family in the TGFβ signaling pathway has previously been described. An initial link between PARPs and components of the TGFβ pathway was proposed by studies with estrogen receptor positive breast cancer cells. The protein PARP1 has been found to interact with a specific promoter region of the gene encoding TβRII and participate to its transcriptional regulation (Sterling et al., 2006). In another study, the involvement of PARylation of Smads as a means for regulating the magnitude and kinetics of TGFβ signaling was characterized for the first time (Lönn et al., 2010). Smad3 and Smad4 are subjected to PARylation by PARP1, a post-translational modification that releases Smads from their target genomic loci and limits Smad-mediated transcriptional responses, as well as physiological responses, such as EMT. The PARylation of Smad3 by PARP1 has also been demonstrated in the context of vascular fibrosis, using rat vascular smooth muscle cells (Huang et al., 2011). Inhibition of the catalytic activity, or down-regulation of PARP1 blocks the nuclear translocation of Smad3. In addition, Smad3 PARylation facilitates the binding of Smad3 to target DNA regions, in order to induce genes related to vascular fibrosis, hindering a positive role of PARP1 on TGFβ-mediated responses. The involvement of PARP1 in the transcriptional
regulation of the genes generating the TβRI and TβRII has been shown in CD⁴⁺ T cells. Inhibition of PARP1 activity enhances the expression of both TβRI and TβRII. In addition, PARP1 binds to TβRII promoter to repress TβRII transcription (Zhang et al., 2013). This study highlights an inhibitory role of PARP1 on TGFβ signaling.

The negative impact of PARP1 on TGFβ signaling has also been shown to occur in an in vivo model for elucidating the effect of PARP1 on prostate tumorigenesis. Genetic ablation of PARP1 in mice, leads to elevated TGFβ, Smad3 and Smad4 levels, with a concomitant induction of EMT markers and down-regulation of epithelial markers. These findings suggest that loss of PARP1 enhances prostate cancer, through TGFβ-mediated EMT (Pu et al., 2014). In a different in vivo model that monitors the effects of the PARP1 inhibitor HYDAMTIQ in the progression of murine lung fibrosis, the TGFβ signaling output is attenuated upon PARP inhibition (Lucarini et al., 2017). In an attempt to unravel the function of PARP1 in the pathogenesis of systemic sclerosis, Zhang et al., have proposed that loss of PARP1 activity results in increased p-Smad3 levels and Smad-dependent transcription of target genes, promoting myofibroblast differentiation in vitro and skin fibrosis in vivo (Zhang et al., 2018c). Thus, PARP1 has a negative impact on TGFβ signaling during skin fibrosis.

The expression of PARP3, a member of the PARP family, is induced in response to TGFβ in different epithelial cell lines and it is positively correlated with mesenchymal phenotypes in breast cancer cells. PARP3 promotes EMT features, such as detachment of neighbouring cells, acquisition of motility, gain of stemness and chemoresistance in breast cancer cells, constituting an important effector of TGFβ-induced EMT (Karicheva et al., 2016).

Overall, the involvement of ADP-ribosylation in TGFβ-mediated responses is of high importance, suggesting that this post-translational modification is as crucial as other well-established modifications (phosphorylation, ubiquitination, and acetylation) for balancing the TGFβ signaling output.
5. Present investigations

TGFβ is a multifunctional growth factor, which initiates a signaling cascade with fundamental roles not only in normal physiological processes but also in pathological conditions. The complicated and often contradictory roles of TGFβ in regulating biological processes, in different cellular contexts, require tight control of its activity. Indeed, cells have developed various mechanisms to modulate the magnitude of the TGFβ pathway. These mechanisms can either positively regulate, and thus strengthen the activity of the pathway, or block the TGFβ signaling cascade, thereby limiting the response of cells to TGFβ. At the molecular level, the regulatory mechanisms, which control the TGFβ pathway, can either target the major components of the pathway, such as TGFβ receptors, and effector Smad proteins, or interfere with downstream Smad-dependent transcriptional responses. In the latter case, the regulatory mechanisms preferentially affect the binding capacity of activated Smads to chromatin or their interplay with co-transcription factors or chromatin remodelers, rather than the initial steps of the TGFβ pathway, i.e. TGFβ receptor activation, R-Smad phosphorylation or nuclear translocation of Smads. Post-translational modifications, such as phosphorylation, ubiquitination and acetylation of members of TGFβ signaling, provide a means of regulation of the pathway. This is achieved, mainly, by altering the stability of Smad proteins or TGFβ receptors or by perturbing the molecular partners of Smads. The TGFβ pathway is also modulated by lncRNAs at multiple levels. They can affect the transcription of genes encoding TGFβ ligands and Smads or they can interfere with the binding of Smads to their target promoter regions, thereby regulating downstream physiological processes.

The aim of our study was to investigate the regulation of TGFβ signaling by lncRNAs and ADP-ribosylation, a post-translational modification, which is mediated by ADP-ribosylating-enzymes and to reveal the molecular basis of their regulatory function. The thesis describes the identification of three relatively unknown lncRNAs, whose expression is regulated by TGFβ signaling. In Paper I, we focus on the molecular characterization of the natural antisense transcript *TGFB2-AS1*, a previously uncharacterized lncRNA. In Paper II we provide, for the first time, evidence that *LINC00707* is a non-coding gene, which responds to TGFβ stimulation. We attempt to unravel the biological significance of this finding. The work presented in paper III demonstrates that the *MIR100HG* lncRNA positively regulates the activity of TGFβ signaling. In Paper IV, we show how PARP1 and PARP2 have a negative impact,
while PARG enhances the TGFβ signaling output. Finally, in Paper V, we show that PARP1, PARP2 and PARG can also control BMP signaling in a similar manner as they do for TGFβ signaling.

**Paper I:** The *TGFB2-AS1* lncRNA regulates TGFβ signaling by modulating corepressor activity

Non-coding RNAs share similar structural properties with mRNAs but they do not generate functional polypeptides. However, they function in the regulation of gene expression and, often, modulate signaling pathways. In this study, we aimed to identify whether TGFβ signaling regulates the expression of genes encoding long non-coding RNAs. For this reason, we profiled the expression of 773 lncRNAs, in response to TGFβ stimulation, using human immortalized keratinocytes. TGFβ signaling regulated the expression of several lncRNAs in a positive or negative way. In addition, some of the TGFβ-target lncRNAs modulated the magnitude of this pathway, suggesting a regulatory feedback loop between TGFβ signaling and a subset of lncRNAs that respond to this growth factor. We focused on elucidating the biological function of the natural antisense lncRNA *TGFB2-AS1*. *TGFB2-AS1* was induced by TGFβ signaling and showed a negative impact on Smad-mediated transcriptional responses. Using an unbiased transcriptomic analysis, we identified several members of the BMP pathway to be down-regulated upon ectopic expression of *TGFB2-AS1*. In contrast, genes involved in the Wnt signaling pathway were up-regulated in *TGFB2-AS1* over-expressing cells, suggesting that this lncRNA can either repress or induce gene expression. In our effort to unravel the molecular mechanism of *TGFB2-AS1* function, we found proteins that interact with *TGFB2-AS1* lncRNA, using mass spectrometry. Interestingly, one of the *TGFB2-AS1*-interacting proteins was the EED protein, a major component of the PRC2 complex. Silencing EED, using RNA interference or inhibition of the PRC2 activity, partially de-repressed the *TGFB2-AS1*-mediated down-regulated genes of the TGFβ and BMP pathways. We conclude that the *TGFB2-AS1*-EED interplay accounts for the inhibitory effects of *TGFB2-AS1* lncRNA on gene expression.

**Paper II:** TGFβ signaling down-regulates *LINC00707* to inhibit inflammatory responses

The human *LINC00707* gene is transcribed to an intergenic lncRNA. *LINC00707* was recently found to act as an oncogene, by promoting cell proliferation and migration in lung adenocarcinoma (Ma et al., 2018). In addition, *LINC00707* induces hepatocellular carcinoma progression, through enhancing
ERK/JNK/Akt signalling, resulting in increased proliferation and colony formation of HCC cells (Wang et al., 2018a). In our study, we describe the transcriptional regulation of LINC00707 by the TGFβ pathway and we attempt to understand the physiological role of this regulation. We observed that TGFβ stimulation led to reduced LINC00707 expression in human immortalized keratinocytes and in human glioblastoma cells. We also demonstrated that LINC00707 is located mainly in the cytoplasm. In order to decipher the biological function of LINC00707, we established keratinocytes where LINC00707 is stably knocked-down and performed RNA-sequencing to identify LINC00707-regulated genes. Comparison of down-regulated genes in response to TGFβ or by LINC000707 silencing revealed genes related to interferon-γ responses, as well as genes involved in chemotaxis of cells of the immune system. On the other hand, the common up-regulated genes between the two aforementioned conditions participate mainly in extracellular matrix organization. In summary, we suggest that TGFβ signaling may establish immune suppressive responses by inhibiting LINC00707 expression.

**Paper III:** The non-coding MIR100HG RNA mediates cytostatic responses of epithelial cells to transforming growth factor β

LncRNAs are classified as intergenic, natural antisense, intronic or bidirectional based on genomic localization criteria. In addition, there are examples of LncRNAs, which serve as primary transcripts that generate mature miRNAs. However, these so-called host genes for miRNAs can also have biological functions, which are inherent to their RNA sequence and do not depend on the miRNA production. This study aimed to identify the biological significance of the regulation of MIR100HG expression by TGFβ signaling. We have demonstrated that TGFβ induced MIR100HG expression in a Smad-dependent way, in normal and cancer cells. In addition, the expression of the miRNAs let-7a-2-3p, miR-125b-5p and miR-125b-1-3p, which are transcribed from the MIR100HG locus were also induced in response to TGFβ. Loss of function experiments showed that MIR100HG silencing compromised the activity of TGFβ signaling and reduced the expression of well known TGFβ-target genes. In contrast, over-expression of MIR100HG resulted in enhanced TGFβ signaling output, a result which was also observed upon over-expression of the mature miRNA let-7a-2-3p. Moreover, ectopic expression of let-7a-2-3p reinforced cell cycle arrest. Lastly, we detected expression of MIR100HG in a panel of different human cancers, suggesting an important, but yet unknown, role of this LncRNA in cancer. In conclusion, we propose that TGFβ-induced MIR100HG expression positively contributes to TGFβ/Smad-mediated transcriptional responses, thus forming a positive regulatory feedback loop.
**Paper IV:** Fine-tuning of Smad protein function by poly(ADP-ribose) polymerases and poly(ADP-ribose) glycohydrolase during transforming growth factor β signalling

ADP-ribosylation is a post-translational modification, which affects molecular interactions and biological processes. The enzymes that catalyze the consecutive addition of ADP-ribose monomers to target proteins (PARPs) play essential roles in diverse cellular processes, such as DNA damage response, gene transcription and apoptosis, through affecting protein-protein or protein-nucleic acids binding. In this paper, we focused on the regulation of TGFβ signaling by members of the PARP family, such as the polymerases PARP1 and PARP-2, and the glycohydrolase PARG. We found that both PARP1 and PARP-2 formed complexes with Smad2/3/4 and ADP-ribosylate them. Silencing of PARP-2 enhanced the TGFβ signaling output, whereas over-expression of PARP-2 has the opposite effect, indicating that PARP-2 plays an inhibitory role in the regulation of TGFβ signaling. On the other hand, PARG also interacted with Smad2/3/4, catalyzed the de-ADP-ribosylation of Smad3 and was required for potent induction of the TGFβ-target genes, thus acting as a positive regulator of the TGFβ pathway. In conclusion, we demonstrated that PARP1 and PARP-2 acted in a similar fashion to negatively regulate TGFβ signaling, whereas PARG quenched the effects of PARPs, thus providing a balance in TGFβ-mediated transcriptional output.

**Paper V:** Regulation of bone morphogenetic protein signaling by ADP-ribosylation

In this work, we report the roles of PARP1 and PARG in controlling the BMP signaling pathway. We show that PARP1 forms complexes with the R-Smads of BMP signaling, Smad1 and Smad5 and induces their ADP-ribosylation, as revealed by in vitro-ADP-ribosylation experiments. BMP signaling output was diminished by PARP1 over-expression, whereas silencing of PARP1 enhanced the expression of the BMP-target genes *ID1* and *SMAD7*. We also demonstrated that the region of Smad1, which is responsible for its association with PARP1, is the MH1 domain, which also serves as the acceptor site for ADP-ribosylation. In particular, single point mutation of lysine 53 to alanine (K53A) in the Smad1-MH1 domain disrupted the ADP-ribosylation of Smad1, without affecting the interaction between PARP1 and Smad1, suggesting that K53 is the acceptor residue of ADP-ribose units. In addition, PARG interacted with Smad5 and Smad4, but not Smad1 and de-ADP-ribosylated Smad1 and Smad5, potentiating BMP signaling. In conclusion, PARG is a positive regulator, while PARP1 is an inhibitory modulator of BMP signaling.
6. Future perspectives

The overall work presented in this thesis highlights the molecular function of three TGFβ-regulated lncRNAs and members of the PARP family. Although significant progress has been made to unravel the molecular function of these factors, understanding their mechanism of action is still in its infancy. There are several questions that remain open and require further investigation in the future.

Paper I: The TGFB2-AS1 IncRNA regulates TGFβ signaling by modulating corepressor activity

In this paper, we suggest that the TGFB2-AS1 IncRNA can alter gene expression profiles. Gene ontology analysis revealed that TGFB2-AS1 negatively regulates genes involved in BMP signaling and a subset of TGFβ-related genes. On the other hand, it up-regulates genes, related to Wnt signaling pathway and many other genes. Our proposed mechanism of the mode of action of TGFB2-AS1 emphasizes the role of the adaptor protein EED as a molecular bridge that mediates the interaction between the PRC2 complex and TGFB2-AS1. Many lncRNAs have been suggested to interact with the PRC2 complex, mainly through the EZH2 or SUZ12 subunits. However, these two proteins do not possess known RNA-binding domains and the level of specificity of their binding to lncRNAs is still under debate (Davidovich and Cech, 2015). We propose that, in the case of TGFB2-AS1 IncRNA, the third core subunit of PRC2 complex, EED, may mediate the IncRNA-PRC2 association. The EED protein contains seven WD40 domains, which allow it to recognize the H3K27me3 modification on chromatin and bind to it. However, WD40 domains can serve as potential RNA-binding surfaces (Jin et al., 2016). Thus, identification of the EED protein domain, which is responsible for this interaction, is necessary. In addition, we suggest that the 3’ end of TGFB2-AS1 mediates its binding to EED. A more thorough mapping of the TGFB2-AS1/EED interaction is needed, in order to narrow down the region of the IncRNA and identify the nucleotides that are absolutely necessary for this binding. Then, point mutations of these nucleotides that abrogate its associa-
tion with EED and functional assays, using mutant TGFB2-ASI should be performed, in order to assess whether the effects of TGFB2-ASI on gene expression are lost in TGFB2-ASI mutants.

The association of TGFB2-ASI with the PRC2 complex can explain the repressive effects of TGFB2-ASI on gene expression. However, over-expression of TGFB2-ASI leads to up-regulation of a substantial number of genes. The positive regulation of gene expression by TGFB2-ASI could be explained by at least two theories. First, TGFB2-ASI may inhibit the expression of master transcriptional repressors that normally maintain the expression of these genes at low levels. This possible function could be mediated by the PRC2 complex. Second, the association of TGFB2-ASI with transcription activators may be hindered, thereby promoting increased binding to promoters/enhancers of target genes and enhanced transcription. Indeed, our mass spectrometry experiment revealed a few members of the Mediator complex (MED1, MED4 and MED21) as protein interactors of TGFB2-ASI. The Mediator complex functions as a general transcriptional activator that positively regulates the expression of several genes, in cooperation with more specific transcription factors. On the other hand, we cannot exclude the possibility that TGFB2-ASI does not facilitate the recruitment of transcription factors to chromatin but rather acts as a molecular decoy which prevents the binding of transcription activators or repressors to chromatin, and thus regulates gene expression in a positive or negative manner. One more question is how TGFB2-ASI can direct the PRC2 complex at specific regions of the chromatin to regulate gene expression. We have demonstrated that TGFB2-ASI can function in trans, as its ectopic expression alters the expression of many genes that are located at regions far from the TGFB2-ASI locus. It would be interesting to elucidate whether this nuclear lncRNA can directly bind chromatin regions, as other lncRNAs do, and transfer the PRC2 complex at specific genomic areas. An exciting idea that requires to be tested is whether TGFβ signaling employs TGFB2-ASI to inhibit genes of the BMP pathway, due to base pairing of TGFB2-ASI with DNA motifs, which are commonly shared between the promoters of BMP-related genes. This possibility could provide an immediate and energetically efficient way for the TGFβ pathway to regulate a subset of genes, without the requirement to produce regulatory proteins that need to be translated and find their proper localization in the genome. Nonetheless, the precise mechanism of the molecular function of TGFB2-ASI needs deeper investigation.
**Paper II:** TGFβ signaling down-regulates LINC00707 to inhibit inflammatory responses

In this work, we identified LINC00707 as a novel TGFβ-target gene. We have demonstrated that LINC00707 is down-regulated by TGFβ signaling but we still need to understand the biological reason of this regulation. We have shown that LINC00707 positively regulates the expression of interferon-γ-related genes, whereas it inhibits some genes involved in the organization of extracellular matrix. The next question, which derives at this stage is how LINC00707 can regulate these groups of genes. One possible idea could be that LINC00707 may regulate a master transcription factor that regulates these processes. This regulation could take place at the transcriptional level or could be a consequence of physical interaction between LINC00707 and this yet unknown protein. Another possibility involves the physical interaction of LINC00707 with mRNAs, which may alter their stability or with miRNAs, thereby sponging them away from their target mRNAs at the cytoplasm.

**Paper III:** The non-coding MIR100HG RNA mediates cytostatic responses of epithelial cells to transforming growth factor β

In this study, we provide evidence that the TGFβ-induced lncRNA MIR100HG acts as a positive regulator of TGFβ signaling. This interesting observation need to be complemented by finding the mode of action of this lncRNA. Our observations support a model whereby MIR100HG could modulate TGFβ signaling independently of the miRNAs, which are transcribed by its third intron. The identification of the subcellular localization of MIR100HG could give useful insight for the molecular function of MIR100HG. In addition, further work is required to identify whether MIR100HG regulates transcription of basic members of the pathway, such as TGFβ ligands, TGFβ receptors and Smads, or interferes with the phosphorylation events that ultimately switch on the signaling. Also, the possibility that MIR100HG affects the binding of Smads to their target promoters needs to be examined. Our work also suggests that TGFβ induces the expression of the miRNA let-7a-2-3p, which has not been described by others. Interestingly, let-7a-2-3p seems to have anti-proliferative properties. These results open new windows for understanding the TGFβ-mediated cell cycle arrest. TGFβ could up-regulate let-7a-2-3p, in order to halt cell proliferation, as a new mechanism, which complements the well-known effects of TGFβ on the transcriptional regulation of the oncogene MYC and the cell cycle inhibitors p21 and p15. Thus, it is interesting to identify the mRNA targets of let-7a-2-3p. Moreover, it would be interesting to investigate whether the lncRNA MIR100HG shows effects, similar to let-
7a-2-3p, on cell proliferation. Concerning the role of MIR100HG on tumor progression and cancer patients' survival, our analysis shows that MIR100HG correlates either with high or with low survival, depending on the type of cancer. Considering the dual role of TGFβ signaling in cancer, we could investigate a potential correlation of the MIR100HG expression to the activity of TGFβ signaling in different tumors and elucidate whether TGFβ is a good or bad prognostic factor for these cancers.

**Paper IV:** Fine-tuning of Smad protein function by poly(ADP-ribose) polymerases and poly(ADP-ribose) glycohydrolase during transforming growth factor β signalling

In this paper we have shown that the PARylation of TGFβ-activated Smads leads to attenuation of TGFβ signaling. In contrast, removal of PARylation from Smad proteins potentiates the signaling. Since PARylation takes place into the nucleus it would be of interest to explore the precise timing of Smad PARylation. For example, PARPs could PARylate activated Smads which are already bound to promoters of target genes, thereby promoting their release from chromatin. Alternatively, PARPs may promote PARylation of Smads and prevent their binding to chromatin regions. In addition, the effects of PARP1, PARP2 or PARG inhibition on TGFβ signaling could be exploited in the treatment of cancer. PARP inhibitors could potentiate TGFβ signaling in types of cancer, which are responsive to tumor suppressor activities of TGFβ. In contrast, PARG inhibitors could be of use during treatment of advanced cancers, in which TGFβ functions as a tumor promoter.

**Paper V:** Regulation of bone morphogenetic protein signaling by ADP-ribosylation

In this paper we propose that PARP1 limits BMP signaling by PARyling Smad1 and Smad5, while PARG removes PARylation from Smads, thereby promoting activation of BMP signaling. In addition, we confirmed that the residue K53 of the MH1 domain of Smad1 is an important acceptor site for ADP-ribosylation. It would be of interest to generate Smad1 and Smad5 mutants, which cannot be ADP-ribosylated and perform functional assays, in order to further check the importance of ADP-ribosylation on Smad1/5-mediated biological outcomes. Moreover, PARP inhibition could be used to evaluate possible effects on BMP signaling in the context of cancer or developmental processes. For example, the cancer model of GBM could be a useful
tool to investigate whether PARP inhibition may have a beneficial function by enhancing BMP signaling, which promotes more differentiated cancer cells and attenuates cell cycle progression in this type of cancer.
Acknowledgments

First of all I would like to thank my main supervisor Prof. Aristidis Moustakas for giving me the opportunity to work in his group and introducing me to the field of TGFβ signaling. Aris, you have been an inspiration for me all these years and one of the main reasons that I decided to continue working in the academia. Your devotion in science and your enthusiasm to explore new paths in our field are a strong motivation for young scientists, who want to follow your paradigm. When I saw you for the first time during one of your seminars in Thessaloniki, back in 2009, I was surprised that I could follow your logic, even though I was ignorant about the topic of your research. However, this detail stimulated me to search about it and to realize that this field is something that I really like. Thus, I believe that you have the charisma to efficiently transmit knowledge to an average audience and eventually motivate young students. I really enjoyed our meetings and the fact that you were always positive to encourage new ideas.

I would also like to thank my co-supervisor Prof. Carl-Henrik Heldin for his guidance, support and useful advice for improvement of the projects, during my PhD studies. Calle, you are probably the most important and at the same time modest person that I have met so far. You are a good example of the rule that says that the most successful and important people are also the simplest and are often among us. I admire your passion about science and your excitement every time we have some promising results to present. I feel lucky that I had you as my co-supervisor.

Moreover, I would like to thank my supervisor during my BSc degree at the University of Thessaloniki, Dr Konstantinos Vlachonasios. During my thesis project at his group I had my first engagement with the laboratory work and I learned the first basic techniques. I would like to thank him for his patience to introduce me to lab work, as well as the Erasmus exchange program.

In addition, I would like to express my acknowledgments to all the past and present members of the TGFβ/STEP group for contributing to an excellent and friendly atmosphere in the laboratory. I would like to start with the person that I worked together, when I started my PhD. Yukihide, you have been an excellent supervisor and colleague. I should especially thank you for your effort to teach me all the basic techniques and make me feel comfortable at the beginning. From the rest of the members of our group, back in 2013, I would like to start with Anita, who is the pillar of the group and the most experienced colleague that I had. Anita, thank you for support all these years. Laia, you
also deserve special acknowledgments for collaborating to my projects, for your useful advice and your spontaneous Spanish/Mediterranean character that I like a lot. Yutaro, thank you very much for teaching me many useful techniques, for discussions concerning experiments, but also non-scientific topics and for collaborating to several projects. Moreover, I would like to thank Varun for working together on the PARP projects. You have been an excellent colleague and most importantly friend all these years. Kallia, you have been one of my best friends in Uppsala and a great co-worker. I enjoyed a lot having you next to me in the lab. Thank you very much for all the funny moments we had and also for your patience to have a messy bench next to yours. Mahsa, thank you for contributing to a nice atmosphere in the lab. You are a very kind and funny person and I enjoyed a lot having you close to me in the lab and sharing your equipment with me sometimes. You have saved me a lot of time when I cannot find my own things on my bench.

I would like to continue with some of the past members of the group, with whom I spent quite a lot of time inside and outside the lab before I go on with the newcomers. Among the old Ludwig colleagues first, I have to thank Jon, who became one of my first friends, when I moved to Uppsala. Jon, you have provided us with very funny moments that I still recall many times. You have definitely contributed to a friendly working environment and I am very lucky to have worked with you. In addition, I would like to thank Erna, Sinnisky, Kaoru and Peter Lönn. I would also like to acknowledge our students Oscar, Simon and Gad, Katia, Natasa, Rafailla and Pedro. From the IMBIM part I would like to thank Claudia of course. Claudia, although you have left the lab some time ago, I still feel your presence among us. You are a very funny girl with a strong character. I miss the moments when you used to welcome me in the lab using the most popular Greek word. They say that the people who insult a lot are also very clever and I think you are a representative example of this theory. I also thank the newcomers Chrysa and Caroline and I wish you best of luck since you are still at the beginning of your adventure. I also ought to express my warm acknowledgments you to the students that I had the privilege to supervise them during their projects in our group: Paris, Paula, Andrew and Eric. Your contribution to the progress of my projects was very important for me and I wish you all the best to your future careers that I am sure they will be full of success.

From the corridor of the Ludwig Institute, I would like to thank Paraskevi Heldin, Ingvar Ferby and Johan Lennartsson, as well as all the members of their groups (Aino, Chun, Merima, Kaustuv, Giulia, Ana-Rosa, Takeshi, Mari, Natalia, Maria, Glenda, Linda, Haisha, Niki, Kehuan). Also, many thanks to people from the other groups: Anders, Oleks, Ihor, Mariya, Anahita, Per, Masato and Ulla. I would also especially thank Noopur for being a nice and cool friend. Noopur, you are a very good scientist but at the same time a very funny person who enjoys the life and make the others laugh. Per, thank you for the philosophical questions and discussions that
made me forget the experiments and think more deeply about the life. Giulia, you have been a real source of positive energy and I really enjoyed seeing you walking rapidly in our corridor to perform your experiments. Also I really feel to thank Carmen for the time we spent together in the lab during the first year at the Institute. In addition, I would like to thank the groups of Anna-Karin Olsson, Staffan Johansson, Lena Kjellen and Jin-Ping Li that they recently joined our corridor.

From the greek community of the Ludwig Institute I especially thank Ria, who was one of the first persons that I met during my first visit at the Institute. Ria, you are the perfect example of an excellent scientist and at the same time a family woman, with an amazing sense of humor. I also thank Maria for contributing to a sweet atmosphere with her smile and for being a great friend, so many years. I should also thank Panagiotis and George, two great friends who spent some time in our lab, which was certainly enough to create tight bonds with each other. You both brought positive energy to us from Greece. I also have to thank another great friend and scientist, in a different field, the field of chemistry. Raffaello, you have been an amazing friend and I really enjoy spending time with you, especially when you try to connect everything around you to chemistry. I am sure that the field of chemistry will gain a lot of things with your presence in the academia. I also thank you for critical reading of the thesis and useful comments for improvement of the text.

Last but not least of the members of the lab, I would like to thank Costas for being one of my best friends, firstly, and an amazing scientist, with whom I shared many useful discussions, secondly. Costa what can I say about you? We have shared so many things together and the only thing that I regret is that we did not work in a common project together. That would be an experience to remember for sure. I am sure we will do it sometime in the future. Good luck with everything my friend and I wish you to have always friendly reviewers.

I also feel grateful to two greek public benefit organizations, the “Bodossaki Foundation” for supporting me financially during my Master studies and the first two years of my PhD studies, and the “Alexander Onassis Foundation” for funding my studies since October 2017. The contribution of these two foundations in the completion of my research projects was crucial. It also improved the quality of our research and helped me to expand my ideas by performing more experiments.

The next person that I really feel grateful to thank is my companion in the life Cristina. You have been supporting my choices from the beginning and you are the main reason for the peace that I find after a tiring day, which gives me the necessary strength to continue. You entered in my life unexpectedly and you turned it to an amazing adventure. I have been taught by you how to enjoy the present and how to deal with unnecessary worries. I am not sure if I had continued to this path if we were not together. Thank you for everything!
The two persons that I really ought to them everything are my parents Evangelos and Eleni. You have been supporting my education since my childhood and gave everything to me in order to go on in the university and eventually to continue with my PhD. This thesis is dedicated to you and I think this fact says everything about how I feel about you. Ευχαριστώ πολύ για όλα!

I cannot exclude of course my two brothers Nikos and Costas for continuously supporting me and for all the moments we share from my childhood since now. Thank you for everything brothers! In addition, I would like to thank my sister in-law Eleni and my little nephew Panos, who, I hope, will read this thesis in a few years. Finally I would like to thank, my beloved aunt Eva and my uncle George, whom I feel like second parents to me and they really supported me a lot during my studies.
References


A doctoral dissertation from the Faculty of Medicine, Uppsala University, is usually a summary of a number of papers. A few copies of the complete dissertation are kept at major Swedish research libraries, while the summary alone is distributed internationally through the series Digital Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine. (Prior to January, 2005, the series was published under the title “Comprehensive Summaries of Uppsala Dissertations from the Faculty of Medicine”.)